744093

Mathematical Simulation of the Effects of Ionizing Radiation on Semiconductors

David P. Kennedy

International Business Machines Corporation
Components Division
East Fishkill Laboratory
Hopewell Junction, New York 12533

Contract No. F 19628-70-C-0098

Project No. 4608
Task No. 460805
Work Unit No. 46080501

FINAL REPORT
1 February 1970 - 31 December 1971
January 31, 1972

Contract Monitor: Bobby L. Buchanan
Solid State Sciences Laboratory

Prepared for

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS 01730

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
International Business Machines Corporation
Components Division, East Fishkill Laboratory
Hopewell Junction, New York 12533

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
MATHEMATICAL SIMULATION OF THE EFFECTS OF IONIZING RADIATION ON SEMICONDUCTORS

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Scientific. Final. 1 February 1970 - 31 December 1971

Approved 12 May 72

5. AUTHOR(S) *(First name, middle initial, last name)*
David P. Kennedy

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 31, 1972	131	49

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
F 19628-70-C-0098	TR 22.1486
b. PROJECT, TASK, WORK UNIT NOS. 4608-05-01	
c. DOD ELEMENT 62403F	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d. DOD SUBELEMENT 634608	AFCRL-72-0257

10 DISTRIBUTION STATEMENT
A - Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Cambridge Research Laboratories (LQ) L. G. Hanscom Field Bedford, Massachusetts 01730

13. ABSTRACT

This report outlines results of two-dimensional mathematical investigations initiated during the present contract period: the insulated gate field-effect transistor, the Schottky barrier field-effect transistor and the bipolar transistor. In addition, information is presented on the mathematical technique used to solve these semiconductor device problems. The second investigation (Schottky barrier field-effect transistor) is complete and has been published in the technical literature. The other two investigations (insulated gate field-effect transistor and the bipolar transistor) remain incomplete, although the material presented here shows that substantial modification is required in the available theory of these semiconductor devices.

A discussion is presented on a newly developed method for numerically solving the two-dimensional ambipolar diffusion equations for holes and electrons in semiconductor material. All numerical computations of this type utilize underrelaxation as a means to attain computational stability, the present discussion outlines the automation of this underrelaxation process whereby all device calculations can be performed on a "hands-off" basis. Another new aspect of this computational method is the adoption of a formulation that provides rapid convergence of the calculated electric current in investigations of semiconductor devices. Resulting from this computational method is an ability to calculate the volt-ampere characteristics of semiconductor devices with a minimum number of simplifying assumptions.

DD FORM 1473

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Two-dimensional mathematical studies						
Insulated gate, field-effect transistor						
Schottky barrier field-effect transistor						
Bipolar transistor						
Ambipolar diffusion equations						

Mathematical Simulation of the Effects of Ionizing Radiation on Semiconductors

David P. Kennedy

International Business Machines Corporation
Components Division
East Fishkill Laboratory
Hopewell Junction, New York 12533

Contract No. F19628-70-C-0098

Project No. 4608
Task No. 460805
Work Unit No. 46080501

FINAL REPORT
1 February 1970 - 31 December 1971
January 31, 1972

Contract Monitor: Bobby L. Buchanan
Solid State Sciences Laboratory

Prepared for

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS 01730

Abstract

IBM Reference No.--TR 22.1486

This report outlines results of two-dimensional mathematical investigations initiated during the present contract period: the insulated gate field-effect transistor, the Schottky barrier field-effect transistor and the bipolar transistor. In addition, information is presented on the mathematical technique used to solve these semiconductor device problems. The second investigation (Schottky barrier field-effect transistor) is complete and has been published in the technical literature. The other two investigations (insulated gate field-effect transistor and the bipolar transistor) remain incomplete, although the material presented here shows that substantial modification is required in the available theory of these semiconductor devices.

A discussion is presented on a newly developed method for numerically solving the two-dimensional ambipolar diffusion equations for holes and electrons in semiconductor material. All numerical computations of this type utilize underrelaxation as a means to attain computational stability; the present discussion outlines the automation of this underrelaxation process whereby all device calculations can be performed on a "hands-off" basis. Another new aspect of this computational method is the adoption of a formulation that provides rapid convergence of the calculated electric current in investigations of semiconductor devices. Resulting from this computational method is an ability to calculate the volt-ampere characteristics of semiconductor devices with a minimum number of simplifying assumptions.

In two spatial dimensions, calculations have been performed to demonstrate that a soft current saturation is characteristic of JFET structures containing a low conductance semiconductor substrate. The calculated volt-ampere characteristics of this device are presented in conjunction with the mobile carrier distribution that results when the structure is biased well into electric current saturation. These investigations include the consequences of a velocity limited transport of carriers within the source-drain channel of a JFET. Resulting from these calculations is a mathematical evaluation of the influence a low conductivity substrate has upon the electrical properties of this semiconductor device; this topic is of particular importance, since the low conductivity substrate material is vulnerable to radiation.

Numerous solutions have been completed for a boundary value problem that closely approximate the two-dimensional cross section of an insulated gate field-effect transistor. These solutions have established the detailed steady-state mechanisms of operation within this semiconductor device. For example, it is shown that channel pinch-off, as outlined in elementary IGFET theory, cannot be verified by a rigorous mathematical analysis. An important result of this effort is a detailed understanding of the physical mechanisms producing electric current saturation in an IGFET, and a recognition that the concept of a well defined threshold voltage cannot be established in the same manner it has been previously established from one-dimensional IGFET theory. These calculations also provide insight into some problems associated with our present concept of electrical conduction in a source-drain channel. Computational experiments have been used to demonstrate basic differences between the physical mechanisms of IGFET operation, as established through a rigorous two-dimensional solution of this problem, and as determined from theoretical concepts presently available in the technical literature.

In the theory of bipolar transistor operation, substantial importance is given to high current h_{FE} fall-off and emitter current crowding. Present theoretical concepts of these observed characteristics are based upon one-dimensional and pseudo two-dimensional solutions of the bipolar transistor; there is little (or no) verification of these solutions, using a rigorous two-dimensional analysis of the problem. From work conducted under this contract, it is shown that substantial h_{FE} fall-off is attributable to two-dimensional mechanisms involved in transistor operation. Furthermore, it is shown that one consequence of these mechanisms is that substantially less emitter current crowding is observed than would be predicted using available bipolar transistor theory. One consequence of this effort is a realization that we must reevaluate many mechanisms that heretofore proposed that the high current fall-off of current gain and f_t is a consequence of current crowding; a detailed understanding of this h_{FE} and f_t problem is not presently available.

Work Statement

With partial support from Air Force Cambridge Research Laboratories under contract F 19628-70-C0098, Project 4608, the International Business Machines Corp., Components Division, E. Fishkill, N.Y. has conducted mathematical investigations directed toward a study of the influence of ionizing radiation upon the operation of semiconductor devices. A principal objective of this effort is to design and develop computational techniques whereby, in two spatial dimensions, the ambipolar diffusion equations for holes and electrons in a semiconductor can be solved for numerous different semiconductor devices. This effort has been undertaken with the realization that a detailed knowledge of semiconductor device operation is an important step toward understanding the consequences of operating these devices in a radiation environment.

In conjunction with the development of these computational techniques, problems of semiconductor device operation have been undertaken under this contract. These problems were selected on the basis of mutual interest to the Air Force and to International Business Machines Corp. By undertaking these problems, new computational techniques have been developed which greatly simplify the process of mathematically analyzing the operation of semiconductor devices. In addition, extensive insight has been gained into the mechanisms of operation within semiconductor devices.

A substantial effort has been directed toward a two-dimensional mathematical analysis of IGFET operation. This analysis was used as a basic vehicle to refine the computational techniques developed under this contract and, in addition, to gain a detailed understanding of the mechanisms involved in the operation of this semiconductor device. A principal goal for this effort was to develop two-dimensional computational methods whereby, the volt-ampere characteristics of this structure can be calculated with sufficient accuracy, and using a reasonable amount of computer time.

In addition to this investigation of IGFET operation, a two-dimensional study was made concerning the physical mechanisms involved in the operation of JFET structures containing a low conductivity substrate. This subject is of importance to a particular class of integrated circuits because it provides substantial insight into the role played by the low conductivity substrate material, which is particularly vulnerable to radiation.

Another application of these computational methods, developed during this period, is a two-dimensional solution of the bipolar transistor. There are two purposes for undertaking the solution of this problem: first, to test the numerous pseudo two-dimensional solutions that are described in the technical literature and, second, to establish that this extremely difficult problem can be solved in a rigorous fashion using presently available high-speed electronic computers. Much of the radiation hardening work under way in this country is based upon theoretical calculations using pseudo two-dimensional techniques, despite the fact that these techniques have not been subjected to a rigorous evaluation.

Contents

	page
Abstract	iii
Work Statement	v
Contents	vii
List of Illustrations	viii
Chapter I Two-Dimensional Analysis of JFET Structures Containing a Substrate	
1.0 Introduction	1
1.1 Mathematical Methods	1
1.2 The High-Low Semiconductor Junction	4
1.3 Calculated JFET Characteristics	10
1.4 References	17
Chapter II The Insulated Gate Field-Effect Transistor	
2.0 Introduction	19
2.1 Analysis	21
2.2 Source-Drain Channel "Pinch-off" in an IGFET	25
2.3 Electric Current Saturation	32
2.4 The Depletion Layer in an IGFET	48
2.5 The Threshold Voltage in an IGFET	62
2.6 The Potential Distribution in an IGFET	64
2.7 References	76
Chapter III Bipolar Transistor Investigations	
3.0 Introduction	77
3.1 Analysis	79
3.2 Current Gain in a Bipolar Transistor	80
3.3 Emitter Current Crowding	93
3.4 References	97
Chapter IV Computational Methods	
4.0 Introduction	99
4.1 Definitions and Notation	101
4.2 Computation Scheme	102
4.3 Relaxation Factors	104
4.4 Difference Approximations	106
4.5 A Critical Factor	112
4.6 Summary - Numerical Experiments	113
4.7 Bipolar Transistor Model	114
4.8 Computation Scheme for Bipolar Transistor Model	117
4.9 References	120

List of Illustrations

Figure	Title	Page
1	Analytical model of a symmetrical JFET	3
2	Analytical model of an asymmetrical JFET containing a low conductivity semiconducter substrate	3
3	Impurity atom and mobile electron distribution in an n^+ - n type of high-low junction	6
4	Calculated majority carrier distribution in the low-doped region of a forward biased high-low junction (silicon)	8
5	Calculated minority carrier distribution in a forward biased high-low junction (silicon)	9
6	Calculated majority carrier distribution in the low-doped region of a reverse biased high-low junction (silicon)	11
7	Calculated minority carrier distribution in a reverse biased high-low junction (silicon)	12
8	Calculated volt-ampere characteristics of the JFET structures shown in Figs. 1 and 2	13
9	Calculated contours of constant majority carrier density within a JFET containing a substrate (constant carrier mobility)	14
10	Calculated contours of constant majority carrier density within a JFET containing a substrate (field dependent carrier mobility)	15
11	Analytical model used in this investigation	24
12	Computational model used in this investigation	24
13	Illustration of the electric field at the semiconductor-oxide interface in an IGFET	26
14	Calculated minority carrier distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 1.0 volt)	28
15	Calculated minority carrier distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 6.0 volts)	29

Figure	Title	Page
16	Calculated mobile carrier flux distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 1.0 volt)	30
17	Calculated mobile carrier flux distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 6.0 volts).	31
18	Calculated volt-ampere characteristics of a long-channel IGFET assuming a constant drift mobility for carriers in the source-drain channel (W_i= 1000 Å, L_c= 10 μm)	34
19	Mathematical model of a wide gate IGFET	37
20	Calculated volt-ampere characteristics for the wide-insulator IGFET shown in Fig. 19 (constant drift mobility)	38
21	Calculated mobile carrier distribution in the source-drain channel of an IGFET containing a wide insulator (Fig. 19)	39
22	Calculated mobile carrier distribution in the source-drain channel of an IGFET containing a wide insulator (Fig. 19)	40
23	Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 10 μm (constant carrier mobility)	42
24	Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 10.0 μm (field dependent carrier mobility)	43
25	Calculated minority carrier distribution in an MOSFET assuming a field dependent carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 6.0 volts)	45
26	Calculated mobilt carrier flux distribution in an MOSFET assuming a field dependent carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 6.0 volts)	46
27	Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 1.0 μm (constant carrier mobility)	47
28	Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 1.0 μm (field dependent carrier mobility)	49
29	Calculated depletion layer edge in a 5.0 μm MOSFET assuming V_{SG}= 4.0 volts and V_{SD}= 5.0 volts	50

Figure	Title	Page
30	Calculated gate electrode space-charge layer width in a 10.0 μm MOSFET when V_{SD}= 0	52
31	Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG}= 2.0 volts and V_{SD}= 0	54
32	Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG}= 2.0 volts and V_{SD}= 0.5 volts	55
33	Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG}= 2.0 volts and V_{SD}= 1.0 volt	56
34	Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG}= 2.0 volts and V_{SD}= 3.0 volts	57
35	Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG}= 2.0 volts and V_{SD}= 5.0 volts	58
36	Calculated source and drain depletion layer edges in a 5.0 μm MOSFET when V_{SG}= 1.0 volt and V_{SD}= 5.0 volts	59
37	Calculated source and drain depletion layer edges in a 5.0 μm MOSFET when V_{SG}= 2.0 volts and V_{SD}= 5.0 volts	60
38	Calculated source and drain depletion layer edges in a 5.0 μm MOSFET when V_{SG}= 4.0 volts and V_{SD}= 5.0 volts	61
39	Comparison between computer calculated and volt-ampere characteristics and eqs. 7 and 9 for the 5.0 μm structure (V_{th} selected to agree in deep saturation)	63
40	Calculated (eq. 1) potential distribution in an MOSFET N_D= 2 x 10^{15} atoms/cm^2; V_{SD}= 1.0 volt; V_{SG}= 3.0 volts	66
41	Calculated (eq. 1) potential distribution in an MOSFET N_D= 2 x 10^{15} atoms/cm^2; V_{SD}= 6.0 volts; V_{SG}= 3.0 volts	66
42	Calculated contours of constant relative electric displacement between the gate electrode and the drain	68
43	Calculated potential distribution along the oxide-silicon interface in a 10.0 μm MOSFET (V_{SG}= 3.0 volts)	69
44	Calculated potential distribution along the oxide-silicon interface in a 5.0 μm MOSFET (V_{SG}= 3.0 volts)	70
45	Calculated potential distribution along the oxide-silicon interface in a 1.0 μm MOSFET (V_{SG}= 3.0 volts)	71

Figure	Title	Page
46	Calculated potential distribution along the oxide-silicon interface in a 10.0 μm MOSFET (V_{SD}= 3.0 volts)	72
47	Calculated potential distribution along the oxide-silicon interface in a 5.0 μm MOSFET (V_{SD}= 3.0 volts)	73
48	Calculated potential distribution along the oxide-silicon interface in a 1.0 μm MOSFET (V_{SD}= 3.0 volts)	74
49	Analytical model for a bipolar transistor	81
50	Calculated current gain for the bipolar transistor shown in Fig. 49	83
51	Illustrative one-dimensional impurity atom profile for a diffused bipolar transistor	84
52	Calculated emitter transport and base region transport efficiencies for a bipolar transistor (Fig. 49)	86
53	Illustration of the calculated active emitter junction area with increasing emitter current	88
54	Illustration of the calculated minority carrier flux distributions in a bipolar transistor	89
55	Illustration of the calculated region of back injection by the emitter junction with increasing emitter current	91
56	Illustration of the calculated back injection of a diffused emitter in a bipolar transistor	92
57	Illustration of the assumed majority carrier flux in a transistor base region (traditional theory)	94
58	Calculated ratio of current density at the center of emitter and at emitter edge	95
59	Illustration of base region majority carrier flux in a bipolar transistor	96
60	Analytical model and boundary assumed conditions for an IGFET	100
61	Orientation of discrete variables	108
62	a) Boundary conditions at source contact	109
	b) Boundary conditions at oxide interface	110
63	Analytical model for a bipolar transistor	115

CHAPTER I

Two-Dimensional Analysis of JFET Structures Containing a Substrate

1.0 Introduction

In a recent publication M. Reiser[1] discussed the results of a two-dimensional calculation of mechanisms involved in the operation of JFET structures containing a low conductivity substrate. His calculations show that the substrate produces a reduction of saturation resistance, when compared with the saturation resistance of devices containing an ideally insulating substrate.[2] This reduction of saturation resistance is attributed to an increased transverse component of the source-drain channel electric field (with an increase of source-drain voltage) that "weakens the diffusion field below the source and gate so that more carriers are pulled into the substrate". This concept is reinforced by illustrations of the calculated increase of mobile carriers within the substrate, when this structure is biased well into electric current saturation.

The purpose of this chapter is to present a more detailed comparison between the physical mechanisms of operation within JFET structures containing an ideally insulating substrate and within structures containing a low conductivity semiconductor substrate. A conclusion derived from this investigation is that carrier accumulation does, indeed, take place within the substrate material of this semiconductor device, and the mechanisms producing accumulation are attributable to the field dependent mobility of carriers within the source-drain channel.

1.1 Mathematical Methods

It has been shown that the hole and electron distributions in semiconductor material are described by the equations[3]

$$\begin{aligned}
&\text{a)} \quad \operatorname{div} \operatorname{grad} \psi = \frac{-q}{\kappa \varepsilon_0} (N - n + p) \\
&\text{b)} \quad \vec{J}_p = -qD_p \operatorname{grad} p - q\mu_p p \operatorname{grad} \psi \\
&\text{c)} \quad \vec{J}_n = qD_n \operatorname{grad} n - q\mu_n n \operatorname{grad} \psi \\
&\text{d)} \quad \operatorname{div} \vec{J}_p = qR_p \\
&\text{e)} \quad \operatorname{div} \vec{J}_n = qR_n \text{ and} \\
&\text{f)} \quad \vec{J}_T = \vec{J}_p + \vec{J}_n
\end{aligned} \tag{1}$$

by assuming no trapping mechanisms within the structure under consideration.

Equation (1a) is Poisson's equation, and it relates the divergence of the electric field ($\vec{E}$ = - grad ψ) to the electrostatic charge distribution arising from both mobile charge carriers (holes p and electrons n) and immobile ionized impurity atoms (N) within the semiconductor lattice.

Equations (1b) and (1c) give the electric current distribution in a semiconductor arising from the transport of mobile holes and electrons. These equations express the dependency of each electric current component ($\vec{J}_p$ and $\vec{J}_n$) upon the concentration gradients of holes and electrons, the mobility of these charge carriers, and the electrostatic potential gradient (electric field) within the semiconductor material.

Equation (1d) and (1e), the continuity equations for holes and electrons in a semiconductor, are based upon an unspecified mechanism for carrier generation and recombination. For simplicity, the present analysis is based upon the assumption of an infinite minority carrier lifetime within the semiconductor material. This simplification is assumed reasonable because the magnitude of recombination and generation encountered in modern silicon devices has little influence upon the important mechanisms contributing to field-effect device operation.

Equation (1f) states that the total electric current density ($\vec{J}_T$) is a vector sum of the electric current densities due to holes ($\vec{J}_p$) and electrons ($\vec{J}_n$).

Finite difference methods are used to solve this system of equations for the boundary value problems used to approximate a JFET. In particular, the current equations (1b and 1c) are transformed into differential equations for electrical flow lines within this semiconductor structure, and these flow line equations are solved (in conjunction with Poisson equation) for a large array of spatial locations. The details of this computational technique are the topic of discussion in a later chapter.

Throughout this investigation two different JFET structures are considered: a JFET in which the junction "pinches-off" against an ideal insulating surface (Fig. 1), and a JFET in which the junction "pinches-off" against a substrate of low-doped semiconductor material (Fig.2).

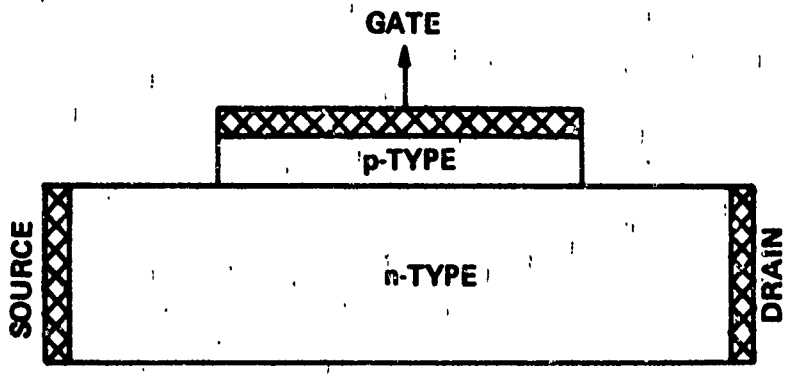

Fig. 1. Analytical model of a symmetrical JFET.

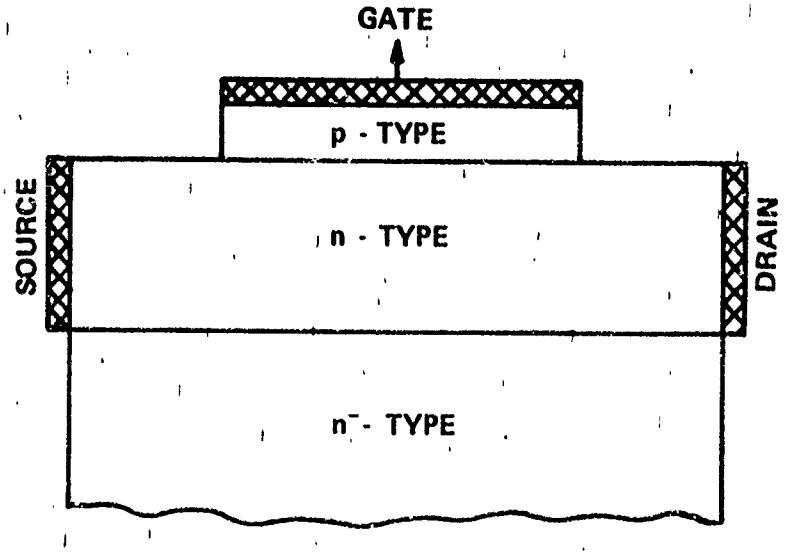

Fig. 2. Analytical model of an asymmetrical JFET containing a low conductivity semiconductor substrate.

The technique for mathematically modeling Fig. 1 was discussed in previous publications on this topic. [2,4] For Fig. 2, a nodal array was selected in which the point density is significantly increased throughout regions requiring detailed mathematical investigation. In particular, an increase of point density is used throughout the boundary between high-doped and low-doped semiconductor material, and throughout that region of the structure in which the gate junction "pinches-off" electric current between the source and drain contacts.

In this analysis particular emphasis is placed upon the selection of boundary conditions that do not introduce errors in the calculated results. For example, the boundary conditions used in this analysis approximate the physical and electrical characteristics of the outer periphery of each semiconductor structure, rather than internal boundaries established by its physical or electrical properties. A homogeneous impurity atom distribution is assumed within each region of Figs. 1 and 2 (gate, channel, substrate), with an abrupt impurity transition at the metallurgical boundary of the gate junctions and at the channel-to-substrate boundary shown in Fig. 2. The exposed semiconductor surfaces are assumed to be an ideal insulator, i.e., no electric current is permitted normal to these bounding surfaces. The ohmic contacts in each model are approximated by equipotantial surfaces that are charge neutral; these ohmic contacts have been located a sufficient distance from the active regions of these devices to have no influence upon their electrical characteristics.

1.2 The High-Low Semiconductor Junction

The channel-to-substrate boundary in Fig. 2 is a high-low semiconductor junction; this topic has been discussed extensively in the technical literature.[4] Because many aspects associated with the operation of this JFET structure (Fig. 2) involve electrical conduction in a high-low junction, a brief review of this subject is presented here. It will be shown that the majority carrier accumulation calculated by M. Reiser[1] is not attributable to conventional mechanisms of electrical conduction across the channel and substrate boundary. Majority carrier accumulation in the low-doped region of a high-low junction is a consequence of conductivity modulation due to the accumulation of minority carriers.

In the calculations of M. Reiser, minority carriers were assumed to have little (or no) influence upon the mechanisms of operation within this semiconductor device; for this reason, the equation for minority carriers was not included in his analysis. Therefore, conventional conductivity modulation (and hence conventional majority carrier accumulation) cannot take place in the mathematical model used by M. Reiser in his publication on this subject.

Figure 3 shows the calculated majority carrier distribution in an n^+-n high-low junction at potential equilibrium. From Fig. 3, an abrupt transition between the high-doped and low-doped material is not representative of the mobile carrier distribution at this boundary. An electrostatic double-layer is formed by a redistribution of majority carriers, relative to the distribution of ionized impurity atoms. Carriers are depleted from the high-doped region thereby forming the positive half of the double-layer; the negative half is formed by these same carriers accumulating on the low-doped side of the high-low boundary. Unlike a p-n junction, the electrostatic charge produced by this double-layer is not significantly influenced by the application of an external biasing voltage. An excessively large biasing voltage (hence, an extremely large current density) would be required to significantly modify this potential barrier without taking into consideration the mechanism of minority carrier conduction.

Electrical conduction across the high-low junction in this JFET involves mechanisms associated with both a forward biased and reverse biased structure. Specifically, if the external biasing voltage is applied in a direction that would tend to reduce the potential barrier in a high-low junction (majority carriers flow from the high-doped to the low-doped material) the structure is presumed to be forward biased. Similarly, if this applied biasing voltage tends to increase this potential barrier (majority carrier flow from the low-doped to the high-doped material) the structure is presumed to be reverse biased. These two different modes of operation have an opposite influence upon the majority carrier density in the low-doped region: forward bias produces carrier accumulation and reverse bias produces carrier depletion.

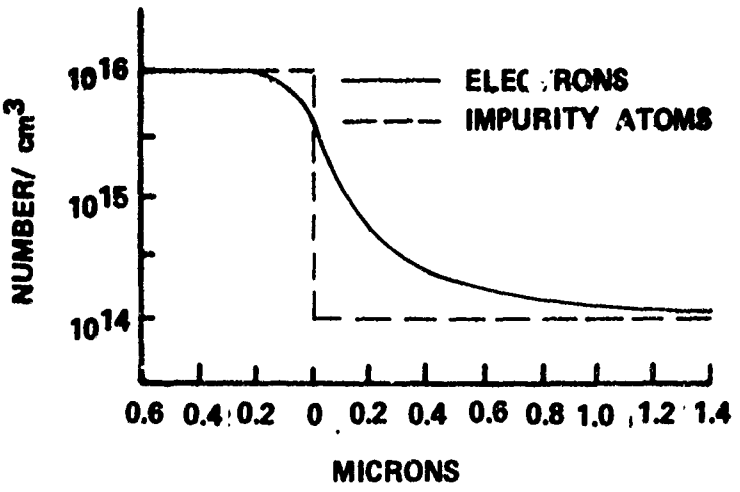

Fig. 3. Impurity atom and mobile electron distribution in an n^+- n type of high-low junction.

In a forward biased high-low junction the electric field distribution at locations removed from the double-layer is determined by the majority carrier distribution and the electric current; this electric field distribution produces electric current continuity, due to majority carriers, throughout this entire semiconductor structure. Majority carrier electric current continuity does not imply a similar situation of electric current continuity for minority carriers, without the intervention of diffusion mechanisms. A detailed analysis of this structure shows that minority carrier transport in the high-doped region is attributable almost entirely to diffusion from the high-low junction to the high-doped region ohmic contact. In contrast, minority carrier current in the low-doped region is attributable to both drift and diffusion. Minority carriers drift toward the high-low junction, although the electric current due to drift far exceeds the electric current due to diffusion on the high-doped side; hence, carrier accumulation takes place in the low-doped material. This process of minority carrier accumulation produces a diffusion gradient in a direction that opposes the drift component of minority carrier current, and thereby current continuity is obtained for minority carriers. This process of minority carrier accumulation will conductivity modulate the low-doped side and produce an accumulation of majority carriers.

This process of carrier accumulation is significantly influenced by the gradient of minority carriers on the high-doped side; an increase of gradient produces an increase of minority carrier current and, hence, a decrease of minority carrier accumulation in the low-doped material. A consequence of this mechanism is a decrease of accumulation as the gate junction "pinches-off" against the high-low junction in Fig. 2, due to an increase of minority carrier current produced by the gate junction space-charge layer.

Figure 4 illustrates the calculated majority carrier accumulation within a forward biased high-low junction. Figure 5 shows the calculated minority carrier distribution producing this majority carrier accumulation (Fig. 4); in addition, Fig. 5 shows minority carrier accumulation in the high-doped side of this structure. Accumulation of minority carriers in the high-doped region is insufficient to produce a significant amount of conductivity modulation.

For a reverse biased high-low junction, the mechanisms producing minority carrier depletion differ little from the mechanisms producing

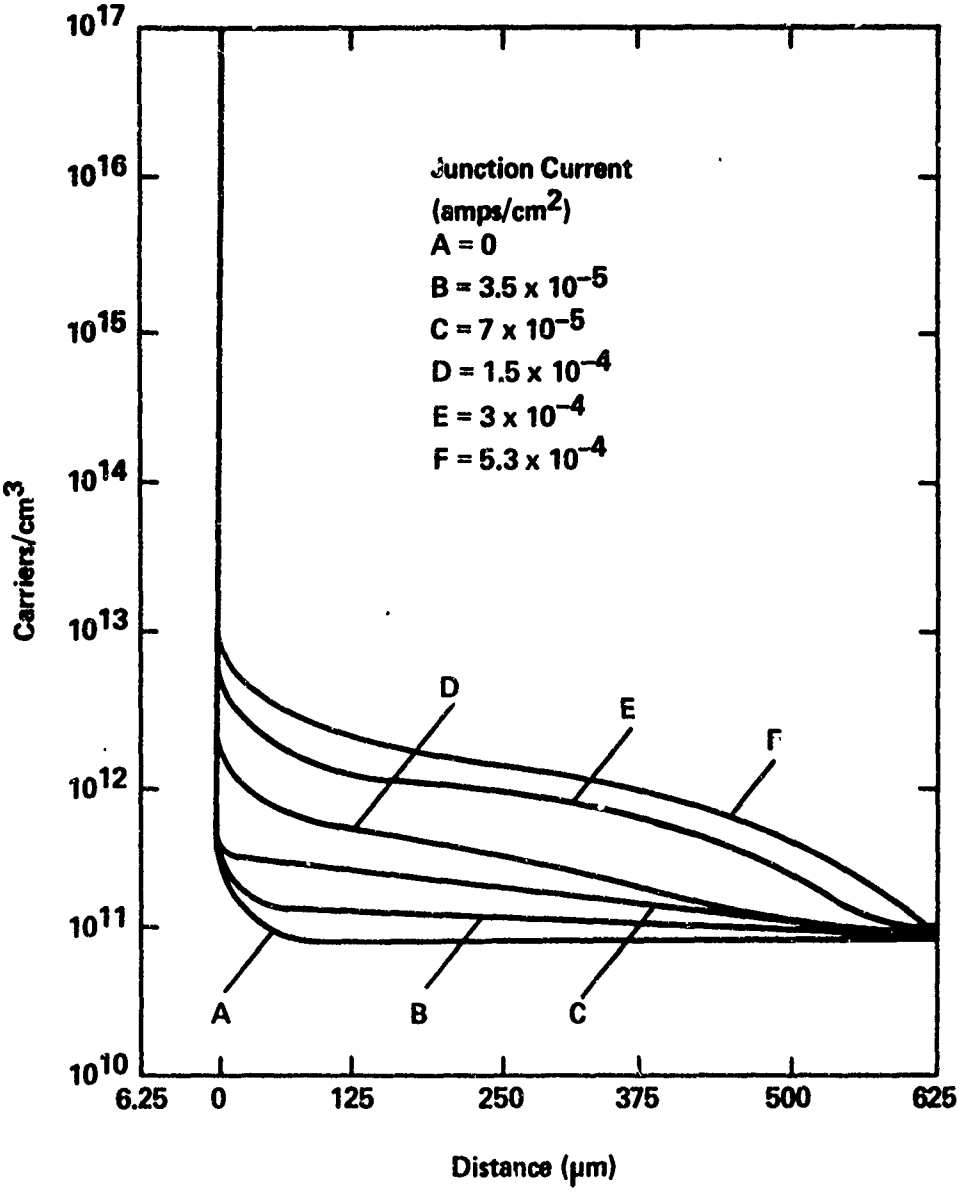

Fig. 4. Calculated majority carrier distribution in the low-doped region of a forward biased high-low junction (silicon).

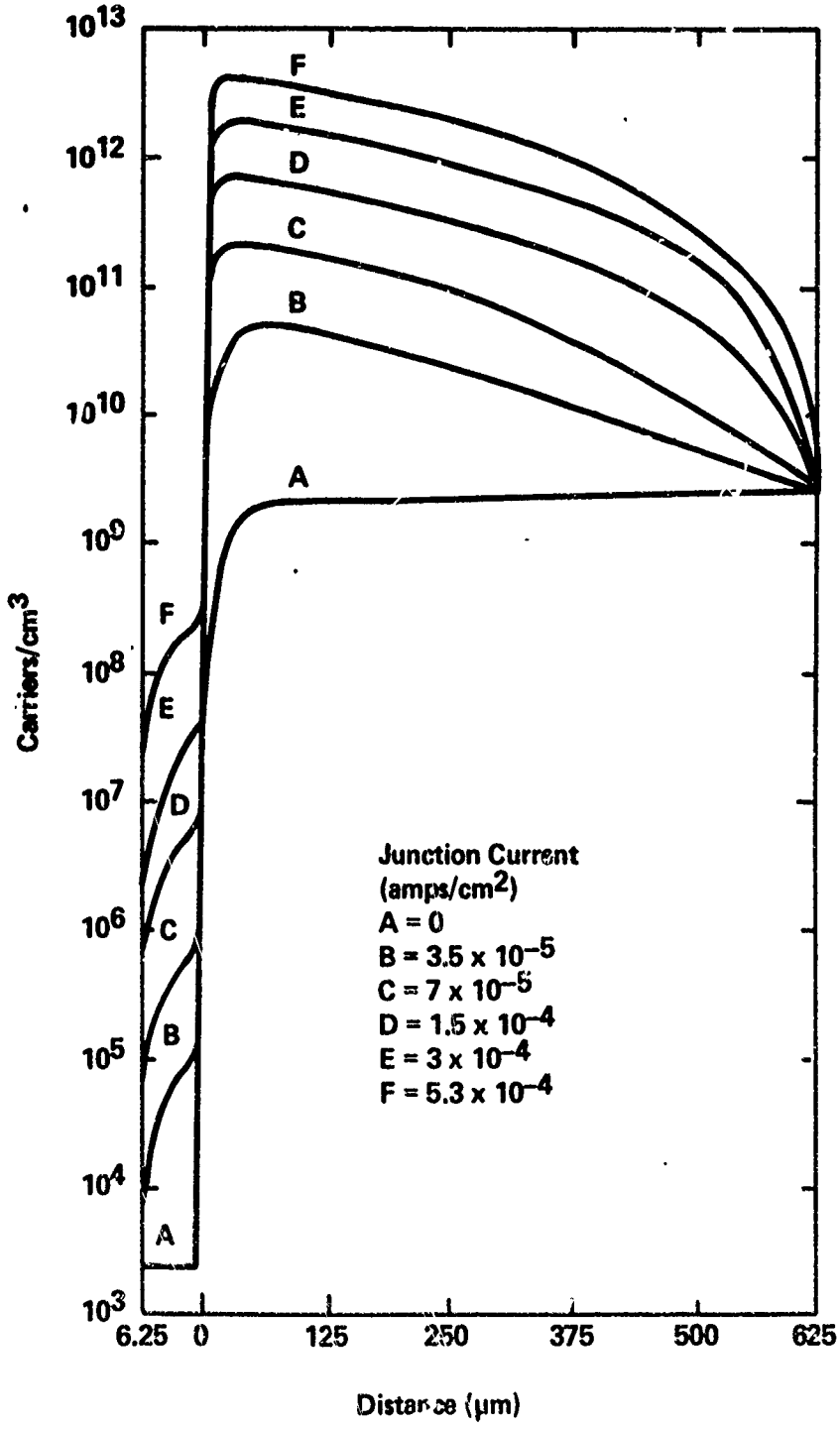

Fig. 5. Calculated minority carrier distribution in a forward biased high-low junction (silicon).

minority carrier accumulation in a forward biased structure. As before, minority carrier current continuity must exist throughout the entire high-low junction. Because the electric field distribution is established by majority carriers, the process of drift alone cannot produce a continuous minority carrier current. For this reason, an additional component of minority carrier current arises through the mechanism of thermal diffusion; in combination, drift and diffusion produce electric current continuity (due to minority carriers) within a reverse biased high-low semiconductor junction.

Figures 6 and 7 show the calculated majority carrier and minority carrier distributions, respectively, in a reverse biased high-low junction. This sequence of calculations demonstrate that minority carrier depletion takes place within both sides of this structure, and that the amount of depletion increases with an increase of electric current. Throughout these regions of carrier depletion charge neutrality is maintained by the removal of an almost equal number of minority and majority carriers, although the change of majority carriers, due to this mechanism, is too small to be observed in Fig. 6.

1.3 Calculated JFET Characteristics

Figure 8 presents a comparison between the calculated volt-ampere characteristics of the two JFET structures shown in Figs. 1 and 2; these calculations are in general agreement with those presented by M. Reiser. In particular, it is shown here that a substantial decrease of saturation resistance (assuming identical channel dimensions and channel doping) can be expected in a JFET containing a low conductivity semiconductor substrate and a field dependent carrier mobility within its source-drain channel. In contrast, Fig. 8 shows this difference of saturation resistance is significantly reduced when a constant carrier mobility is assumed in these structures. From Fig. 8, it is clear that this reduction of saturation resistance is attributable to the low conductivity semiconductor substrate, in conjunction with the assumption of a field dependent carrier mobility.

Figures 9 and 10 illustrate the calculated mobile carrier distribution within this JFET containing a substrate (Fig. 2), when the structure is

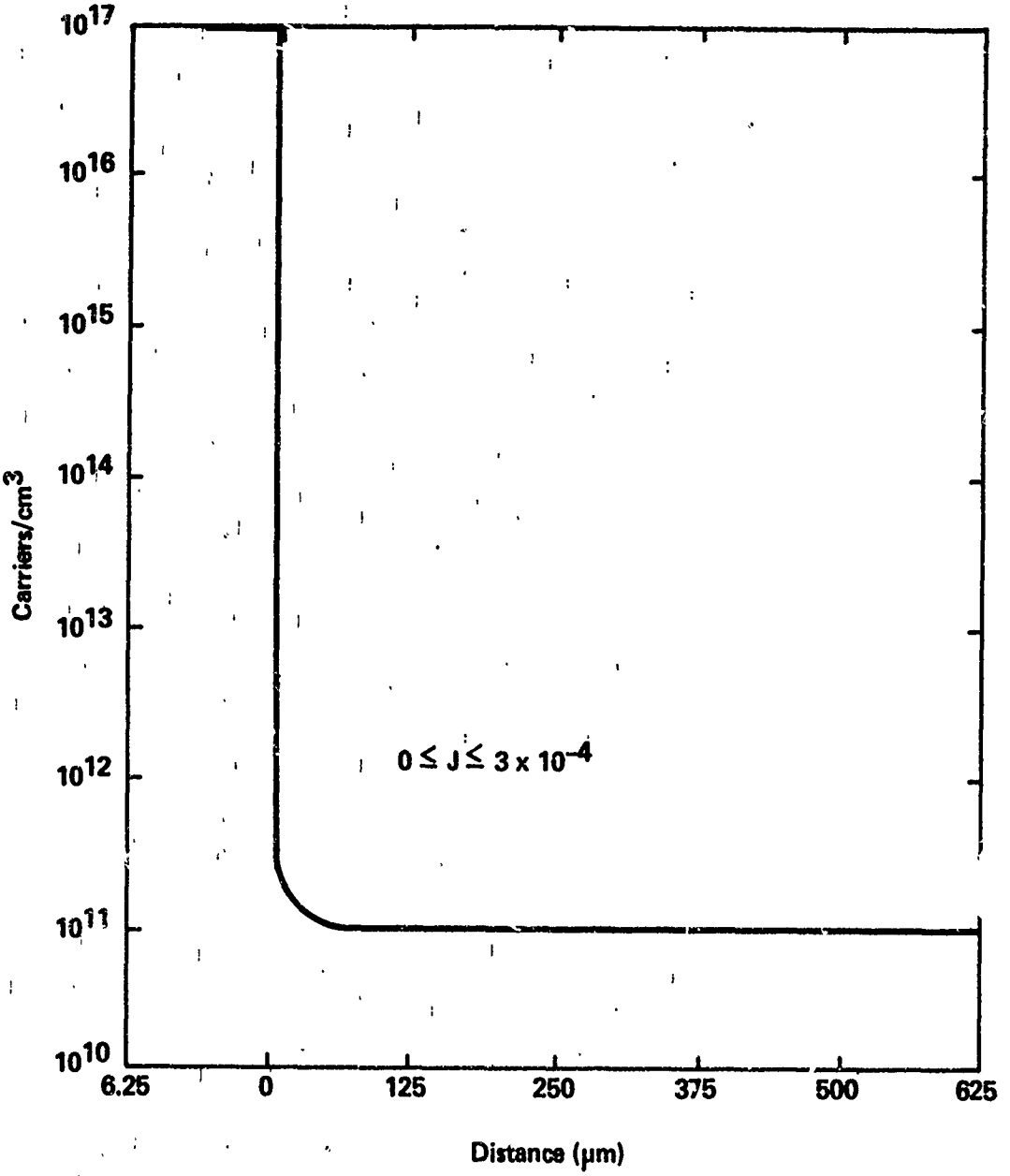

Fig. 6. Calculated majority carrier distribution in the low-doped region of a reverse biased high-low junction (silicon).

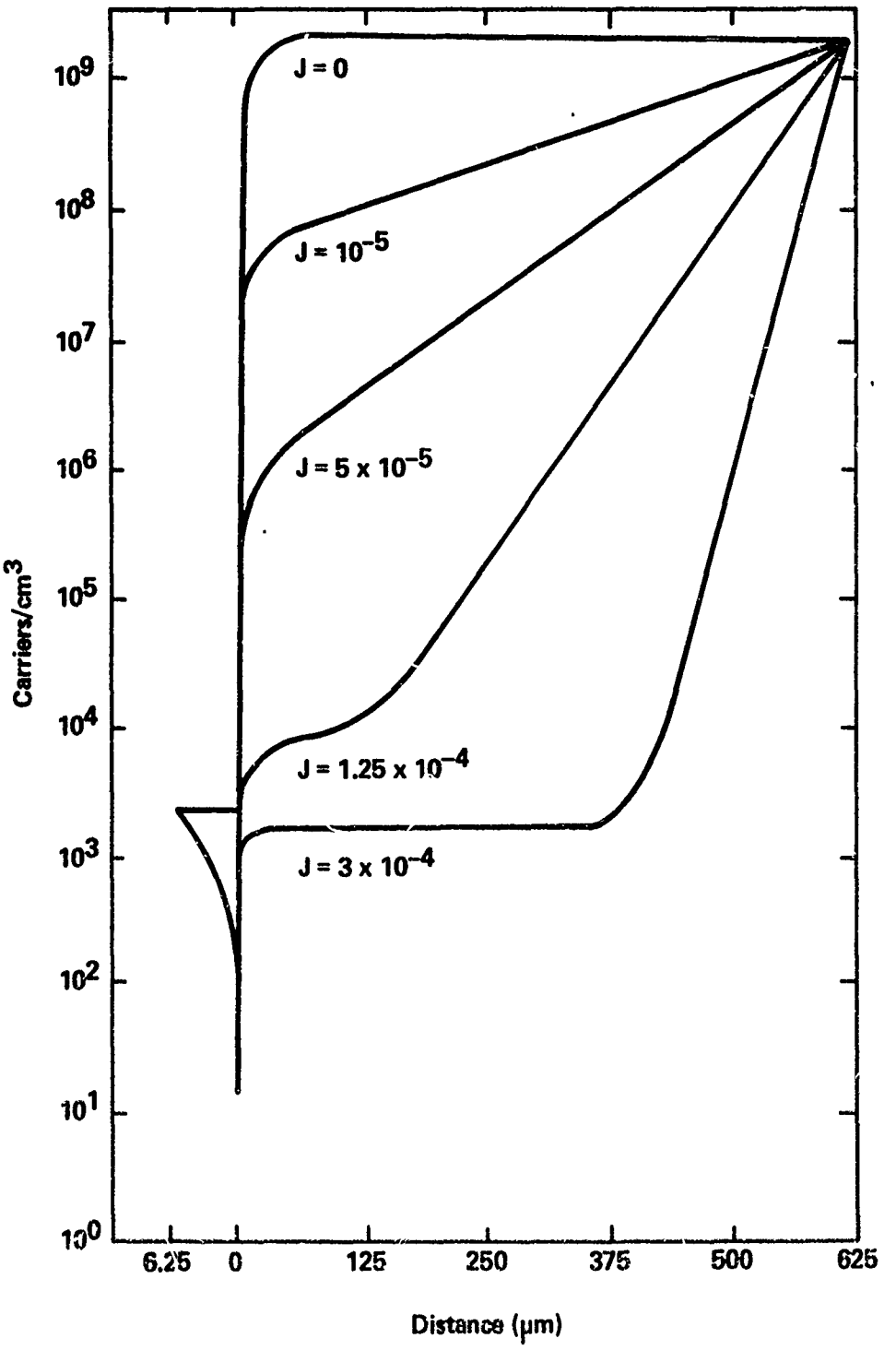

Fig. 7. Calculated minority carrier distribution in a reverse biased high-low junction (silicon).

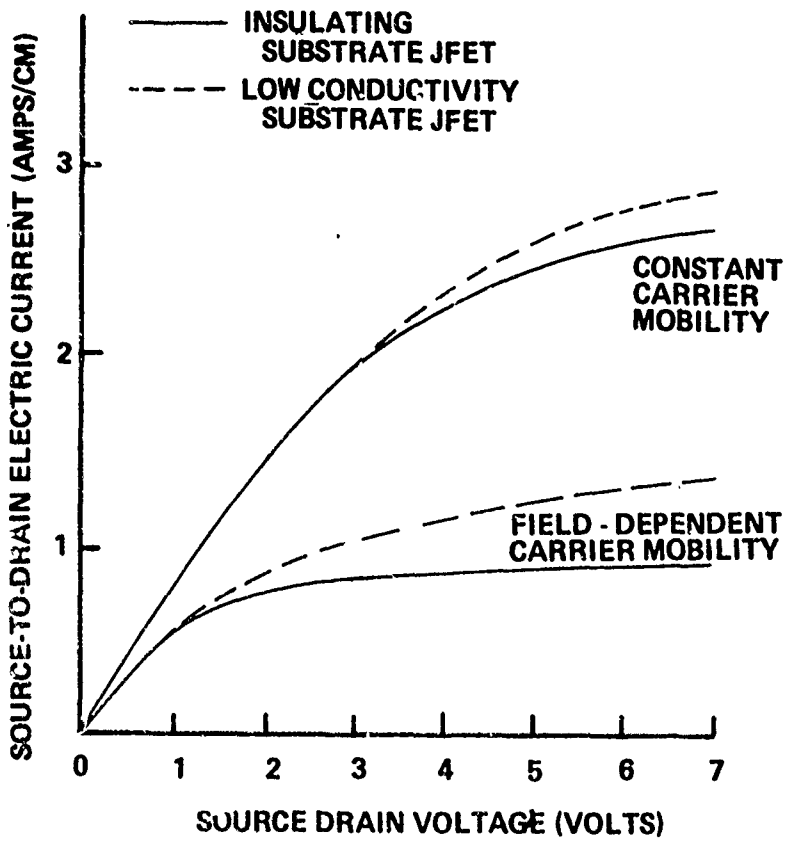

Fig. 8. Calculated volt-ampere charcteristics of the JFET structures shown in Figs. 1 and 2.

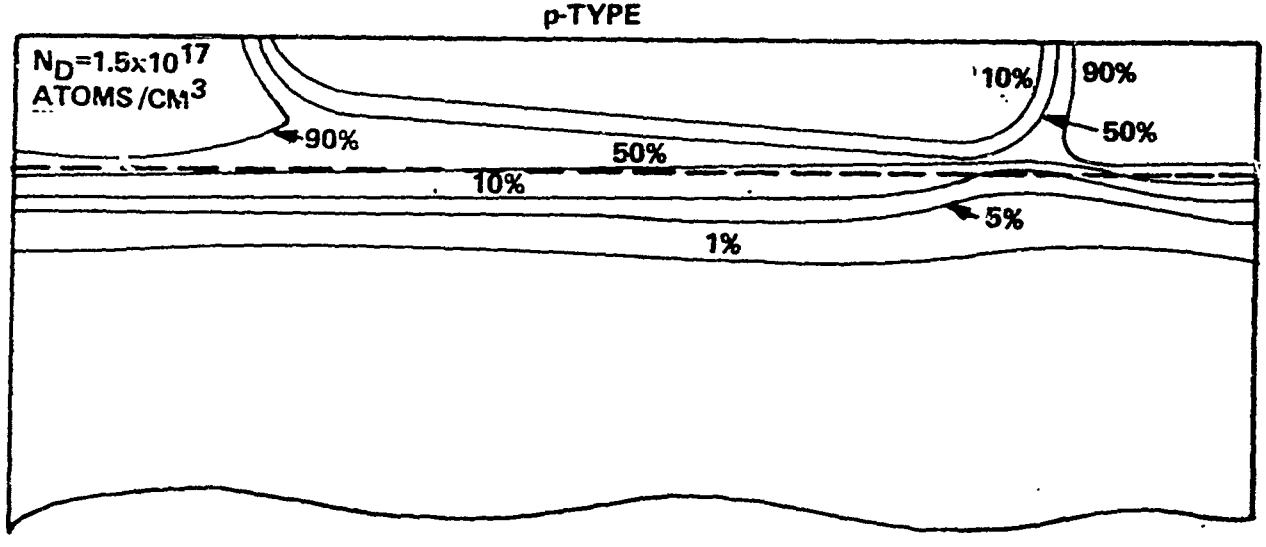

Fig. 9. Calculated contours of constant majority carrier density within a JFET containing a substrate (constant carrier mobility).

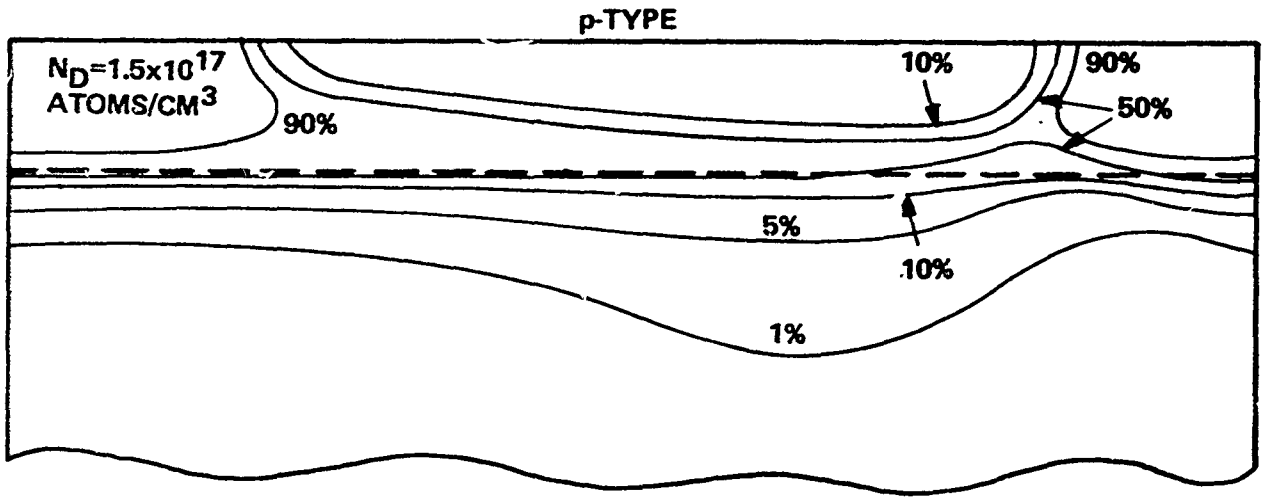

Fig. 10. Calculated contours of constant majority carrier density within a JFET containing a su[illegible] (field dependent carrier mobility).

biased well into current saturation. Shown in this illustration are contours of constant electron density as measured with respect to the impurity atom density in the source-drain channel (1.5×10^{17} atoms/cm^3); for example, the 10% contour represents a mobile electron of 1.5×10^{16} electrons/cm^3. Figure 9 is based upon an assumption that mobile carriers exhibit a constant mobility throughout the entire structure. Figure 10 shows the calculated mobile electron distribution when these carriers are assumed to exhibit a field dependent drift mobility.

From Fig. 9, a constant carrier mobility within the source-drain channel produces a situation where the substrate carrier distribution differs little from the distribution implied by a one-dimensional analysis of high-low junctions (Fig. 3). A layer of excess mobile carriers lies essentially parallel to the metallurgical high-low junction, and this layer exhibits a small contraction near the drain side of the structure. A detailed calculation of the electric current distribution near the "pinched-off" region shows that a substantial part of the total source-drain current arises from carrier transport through this layer of carrier accumulation. For all practical purposes, the effective channel width is increased by this high conductivity layer residing at the channel-to-substrate boundary.

In a previous investigation[2,4] it was shown that electric current continuity is maintained in a JFET (Fig. 1) by the process of carrier accumulation. If the total electric current entering the "pinched-off" region exceeds the total electric current within this "pinched-off" region (due to velocity limited carrier transport) carrier accumulation will take place, and this accumulation provides the means whereby electric current continuity is maintained between the source and drain. Furthermore, it was shown that this region of carrier accumulation will physically expand the conduction path through this "pinched-off" region, and even produce contraction of the gate junction space-charge layer. It is suggested here that the calculated substrate carrier accumulation shown in Fig. 10 (and also in the publication of M. Reiser[1]) is another manifestation of accumulation region expansion.

For this reason, the soft saturation characteristics attributed to the substrate in a JFET (Fig. 2) appears to result from an additional degree of freedom for carrier accumulation within the source-drain channel (when compared with the structure shown in Fig. 1). In the absence of a substrate, carrier accumulation is the dominant mechanism whereby current continuity is maintained; expansion of this region of accumulation is limited by tne gate junction space-charge layer and the insulating surface. In contrast, structures containing a substrate provide semiconductor material in which this region of excess carriers can expand and thereby produce a soft current saturation characteristic.

1.4 References

1. M. Reiser, *Electronics Letters*, 6, 493 (1970)
2. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. & Dev.*, 14, 95 (1970)
3. W. Van Roosbroeck, *Bell Syst. Tech. J.*, 29, 560 (1950)
4. D. P. Kennedy, *Mathematical Simulation of the Effects of Ionizing Radiation on Semiconductors*, Final Report, AFCRL Contract No. F 19628-67-C0111, Feb. 1970

CHAPTER II

The Insulated Gate Field-Effect Transistor

2.0 Introduction

Early investigations of IGFET operation[1] suggest that electric current saturation is attributable to source-drain channel pinch-off in a manner similar to that proposed by Shockley in his publication on the unipolar transistor.[2] Briefly, it was suggested that current saturation is a consequence of the potential distribution between two regions within the source drain channel:

Region I -- from the source junction to the channel pinch-off point, and

Region II -- from the channel pinch-off point to the drain junction.

This concept of IGFET operation has been elaborated on by many authors, yet there appears to be little understanding concerning the detailed mechanisms contributing to this potential distribution.

Another viewpoint concerning electric current saturation is that channel pinch-off produces a region of small electrical conductivity within the source-drain channel, and this region of small conductivity produces the observed saturation resistance.[3] Detailed mechanisms involved in channel pinch-off have been studied by many workers. For example, Chiu and Sah[4] present a two-dimensional analytical solution for the potential distribution throughout this region; due to inherent limitations of their analysis it provides little insight into the solution of this problem. Further, Pao and Sah[5] view channel pinch-off as a "bottleneck" in which diffusion becomes an important mechanism of carrier transport. Two-dimensional numerical solutions of IGFET operation were presented by M. B. Barron[6] and by Vandorpe and Xuong;[7] both of these numerical solutions show the existence of a well defined channel pinch-off point, althouth it has been established that this result is a consequence of computational errors associated with their numerical method[8] (see Chapter IV).

The purpose of this chapter is to present the initial results of a detailed two-dimensional steady-state mathematical investigation of

IGFET operation. The system of equations applied to this problem were previously outlined by Van Roosbroeck[9] in connection with his studies of the transport of electrons and holes in a semiconductor. Finite difference methods are used, with the aid of an electronic computer. Using this technique, a rigorous analysis is under way with an aim toward understanding the physical mechanisms of operation that produce the electrical characteristics exhibited by an insulated-gate field-effect transistor.

This mathematical technique introduces many difficulties of interpretation. The system of equations used in our two-dimensional computer solutions of the IGFET problem are exceedingly general, and implicitly contain all mechanisms important to the operation of this semiconductor device. As a consequence, it is sometimes difficult to explain the results of these computer calculations in terms of recognized physical mechanisms involved in the operation of an IGFET. One method for solving this difficulty is to quantitatively and qualitatively compare the results of our computer calculations with elementary theory taken from textbooks and technical papers; this elementary theory is based upon specific physical mechanisms of device operation. If the computer calculations are in agreement with elementary theory it is presumed the dominant physical mechanisms are known. If, instead, computer calculations disagree with elementary theory, it is presumed that the elementary theory is incomplete; seldom can it be shown that elementary theory is incorrect.

After establishing that a particular aspect of the existing elementary theory is inadequate, an important technique for determining the mechanisms involved in semiconductor device operation (and a technique used extensively in our investigations) is to conduct experiments on the computer. These experiments may consist of new or unusual geometrical configurations, biasing conditions, etc.; each experiment designed to simplify the problem of interpretation. In this fashion, we can often determine the individual mechanisms involved in IGFET operation and, hopefully, put together an elementary theory that is in substantial agreement with the rigorous computer calculations. The overall purpose of this effort is to evaluate existing theoretical concepts of IGFET operation and, where

necessary, modify and improve these concepts to yield an elementary (or engineering) theory that can be used without a computer.

2.1 Analysis

It has been shown that the hole and electron distributions in semiconductor material are described by the equations[9]

$$\text{a)} \quad \text{div grad } \psi = \frac{-q}{\kappa\varepsilon_0}(N - n + p)$$

$$\text{b)} \quad \vec{J}_p = -qD_p \text{ grad } p - q\mu_p p \text{ grad } \psi$$

$$\text{c)} \quad \vec{J}_n = qD_n \text{ grad } n - q\mu_n n \text{ grad } \psi \qquad (1)$$

$$\text{d)} \quad \text{div } \vec{J}_p = qR_p$$

$$\text{e)} \quad \text{div } \vec{J}_n = qR_n \quad \text{and}$$

$$\text{f)} \quad \vec{J}_T = \vec{J}_p + \vec{J}_n$$

when it is assumed that no trapping mechanisms exist within the structures under consideration. Finite difference methods are used to numerically solve this system of equations; details of these computational methods are outlined in Chapter IV.

As in all mathematical investigations of this type there are important physical parameters associated with IGFET operation for which little (or no) information is available. For example, little quantitative date is available concerning either the small field or large field mobility of carriers in a source-drain channel. Similarly, minority carrier lifetime is seldom known with any degree of accuracy. A consequence of this situation is the necessity to make assumptions concerning the physical parameters of an IGFET that may not be completely correct. Throughout this mathematical study we have adopted values for specific physical parameters that are consistent with information presently available in the technical literature. In most situations the quantitative values adopted for these parameters have little infuence upon the physical mechanisms under investigation; when this influence is significant, the topic is discussed in detail.

It has been suggested that the large field transport characteristics of mobile carriers within a source-drain channel exhibit hot electron characteristics that differ from those encountered in bulk semiconductor material.[10] This suggestion appears to be confirmed by recent publications on this topic.[11-13] From these publications, we have assumed in the present analysis that mobile carriers exhibit a maximum (or terminal) velocity of about 8.6×10^6cm/sec; this magnitude is taken from measurements upon bulk silicon,[14] and it is within a factor of 2 of the measured terminal velocity for carriers within an MOS channel.[13] From these publications we have also used traditional assumptions for the warm electron region of the carrier velocity:[12,14]

$$v \alpha E \quad \text{when } 0 \leq v < v_c$$

$$v \alpha E^{1/2} \text{ when } v_c < v < v_m ,$$

where E represents the electric field and the symbols v, v_c and v_m represent the carrier velocity, the critical carrier velocity producing a relation $v \alpha E^{1/2}$, and the maximum (or terminal) carrier velocity, respectively.

The magnitude of electric field producing this critical carrier velocity (v_c) remains an unanswered question. It was recently suggested[11,12] that an electric field of 1.5×10^3cm/sec produces warm electrons ($v \alpha E^{1/2}$) in an MOS channel exhibiting a low field channel mobility of about 250 cm^2/volt sec. Therefore, we have adopted this value of critical carrier velocity in the present two-dimensional analysis, and the calculated results based upon this assumption appear to be in agreement with qualitative measurements upon laboratory MOSFET structures.

Our present method for solving Eqs. 1 is based upon an assumption that no carrier recombination exists within the structure under consideration. It is believed the small recombination/generation rate present within modern silicon IGFET structures has little influence upon its steady-state electrical characteristics. Because the present investigation is restricted to the steady-state properties of this

semiconductor device, there is no reason to suspect that by neglecting recombination/generation in our mathematical model we have introduced a significant computational error.

Throughout this mathematical investigation, particular emphasis is placed upon the selection of boundary conditions that do not introduce errors in the calculated results. In particular, the boundary conditions used in this analysis approximate physical and electrical characteristics at the outer periphery of this semiconductor structure, rather than along internal boundaries established by its physical or electrical properties. A homogeneous impurity atom distribution is assumed within both the n-type and p-type semiconductor material. The exposed semiconductor surface is assumed to be an ideal electrical insulator; i.e., no electric current is permitted normal to these bounding surfaces. The ohmic contacts are approximated by equipotential surfaces that are charge neutral; these ohmic contacts have been located sufficiently far from the active regions of this device to have no influence upon its electrical properties.

In the calculations presented here, the physical and geometrical properties of our analytical model (Fig. 11) were selected as representative of a typical MOS field-effect transistor. All calculations are based upon analytical models that have one of three different distances between the source and drain metallurgical junctions: 10 μm, 5.0 μm, and 1.0 μm. Rather than calculate the source and gate junction shapes shown in Fig. 11, the simplified model shown in Fig. 12 is used to approximate an IGFET; geometrical differences between Figs. 11 and 12 have little influence upon the conclusions presented here.

In the source and drain regions of these structures a donor atom density of 10^{19} atoms/cm^3 is assumed to form ideally abrupt p-n junctions into a substrate region containing an acceptor atom density appropriate to the channel length:

Channel length (μm)	Substrate doping (atoms/cm^3)
10	2×10^{15}
5.0	2×10^{15}
1.0	2×10^{16} .

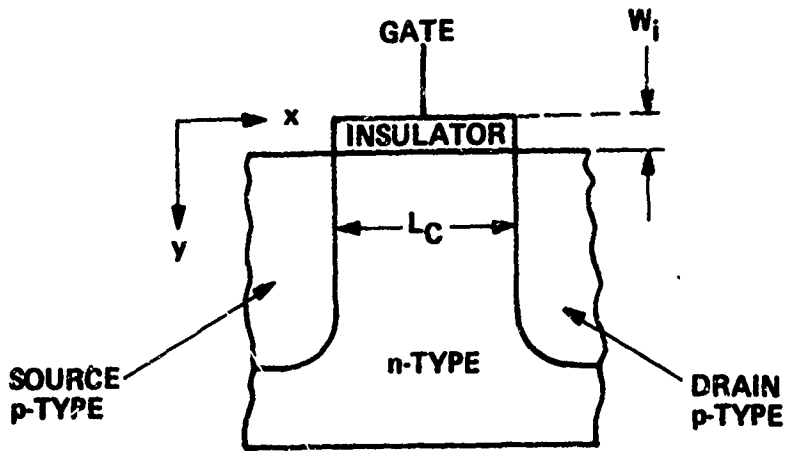

Fig. 11. Analytical model used in this investigation.

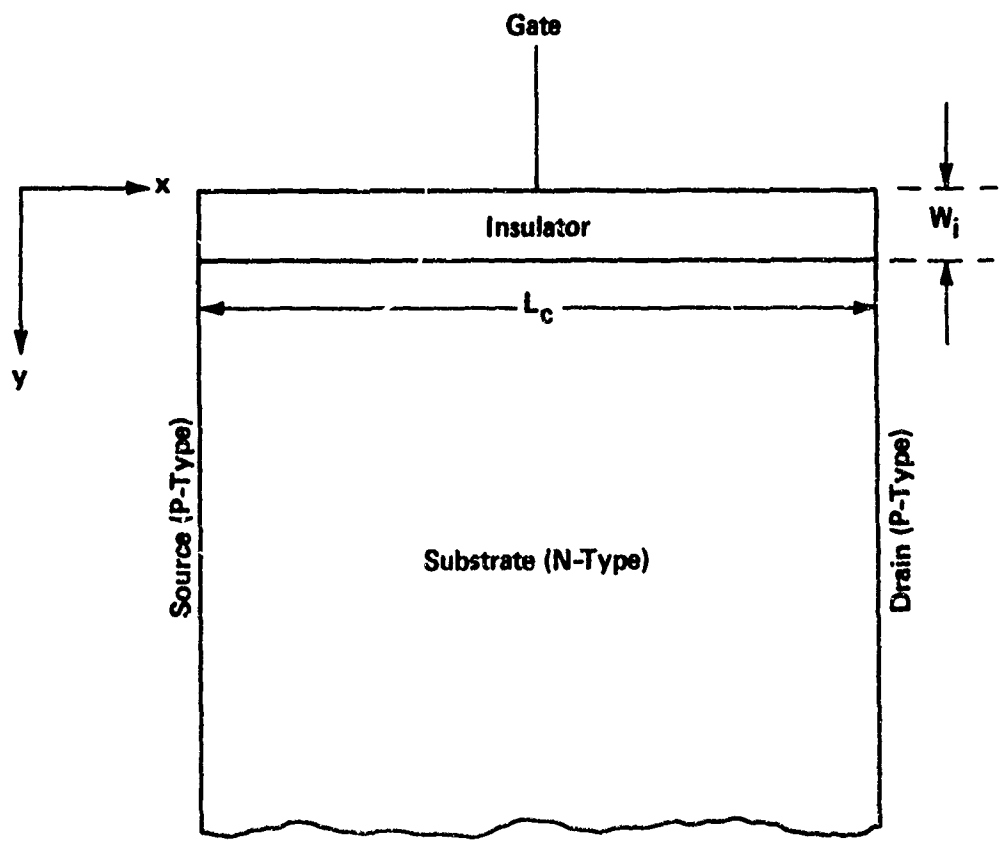

Fig. 12. Computational model used in this investigation.

To obtain a reasonable comparison between the electrical properties of these analytical models, it was necessary to introduce modifications to compensate for changes of substrate doping, and thereby obtain a consistent threshold voltage. For this reason, an oxide (SiO_2) thickness of 1000 Å is assumed in conjunction with the substrate doping of 2×10^{15} atoms/cm^3 and 288 Å with the doping of 2×10^{16} atoms/cm^3.

2.2 Source-Drain Channel "Pinch-off" in an IGFET

In his investigations of JFET operation, Shockley[2] showed that channel pinch-off by the gate junctions produces a region of carrier depletion across which appears that portion of the source-drain voltage in excess of the channel pinch-off voltage. Because current continuity must be maintained throughout this semiconductor structure, it is clear that channel pinch-off in a JFET cannot produce complete depletion of mobile carriers;[15] a sufficient carrier density will always exist throughout the pinch-off region to maintain the source-drain electric current.[16] A similar statement can be made for the pinch-off region of an IGFET if, indeed, the mechanisms of operation in an IGFET are equivalent to those encountered in a JFET. The purpose of this section is to consider the physical mechanisms encountered in the source-drain channel of an IGFET when it is assumed that the structure is biased well into electric current saturation.

For purposes of simplicity, it is assumed here that the source-drain channel pinch-off point in an IGFET is coincident with that location where the difference of potential between the semiconductor-insulator interface and the gate electrode just equals the threshold voltage. A short distance from this location the gate electrode and the semiconductor are at the same potential and no electric field exists in a direction perpendicular to the semiconductor surface, Fig. 13. On the source side of this location the gate induced electric field produces an accumulation of minority carriers along the semiconductor surface. In contrast, near the drain side of this location the gate induced electric field forces mobile carriers away from the semiconductor surface; therefore, between the channel pinch-off point (approximately) and the drain the semiconductor surface is essentially free of mobile carriers.

This situation has a significant influence upon the mechanism of channel pinch-off. Between the drain junction and the channel pinch-off point

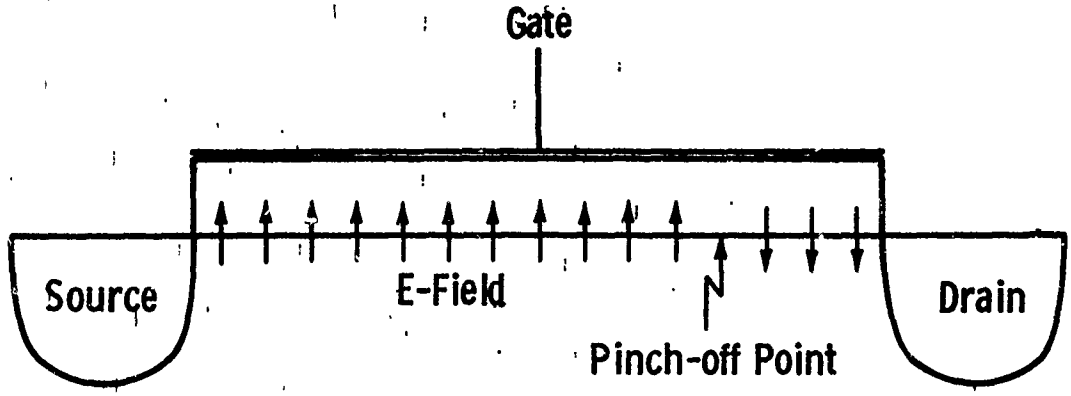

Fig. 13. Illustration of the electric field at the semiconductor-oxide interface in an IGFET.

the difference of potential is established almost entirely by ionized impurities within the drain junction space-charge layer. Further because this difference of potential is always smaller than the applied drain-source voltage, the pinch-off point must reside inside this drain junction space-charge layer, a region of large longitudinal electric field.

If, indeed, channel pinch-off produces a "bottle-neck" mechanism for the source-drain electric current (similar to that encountered in a JFET) a constraint must exist upon the physical width of this conducting channel. In a JFET this constraint is a consequence of an electric field arising from the gate junction space-charge layers.[16] In contrast, the channel pinch-off point in an IGFET is located within a region of large longitudinal electric field; this field provides a mechanism whereby carriers can (and do) leave the channel and continue toward the drain junction. For this reason, pinch-off of the source-drain channel does not produce a constriction of conducting carriers at the semiconductor surface but, instead, it produces an increase in the width of this conducting channel. It is therefore suggested that pinch-off of mobile carriers will not exist in a structure biased well into electric current saturation. This qualitative argument has been verified by a rigorous numerical calculation of the mobile carrier and electric flux distributions in an IGFET.

To illustrate this situation, Figs. 14 and 15 show the calculated minority carrier distributions in an IGFET biased below saturation and in one biased well into current saturation, respectively. Below saturation (Fig. 14) minority carriers form a definite channel between the source and drain. In contrast, when this structure is biased into current saturation (Fig. 15) mobile carriers start to move away from the semiconductor surface at a location far removed from what is normally considered the pinch-off point.

From Fig. 15, it is difficult to maintain the concept of channel pinch-off in an IGFET. This situation is further emphasized in the calculated mobile carrier flux distribution shown in Figs. 16 and 17. Figure 16 shows the flux distribution in an IGFET biased well below current saturation. Similarly, Fig. 17 shows the flux distribution when this structure is biased well into current saturation. Clearly, little is found to justify the existence of pinch-off condition in the theory of IGFET operation (see Fig. 17). Therefore, throughout this report the traditional concept of a

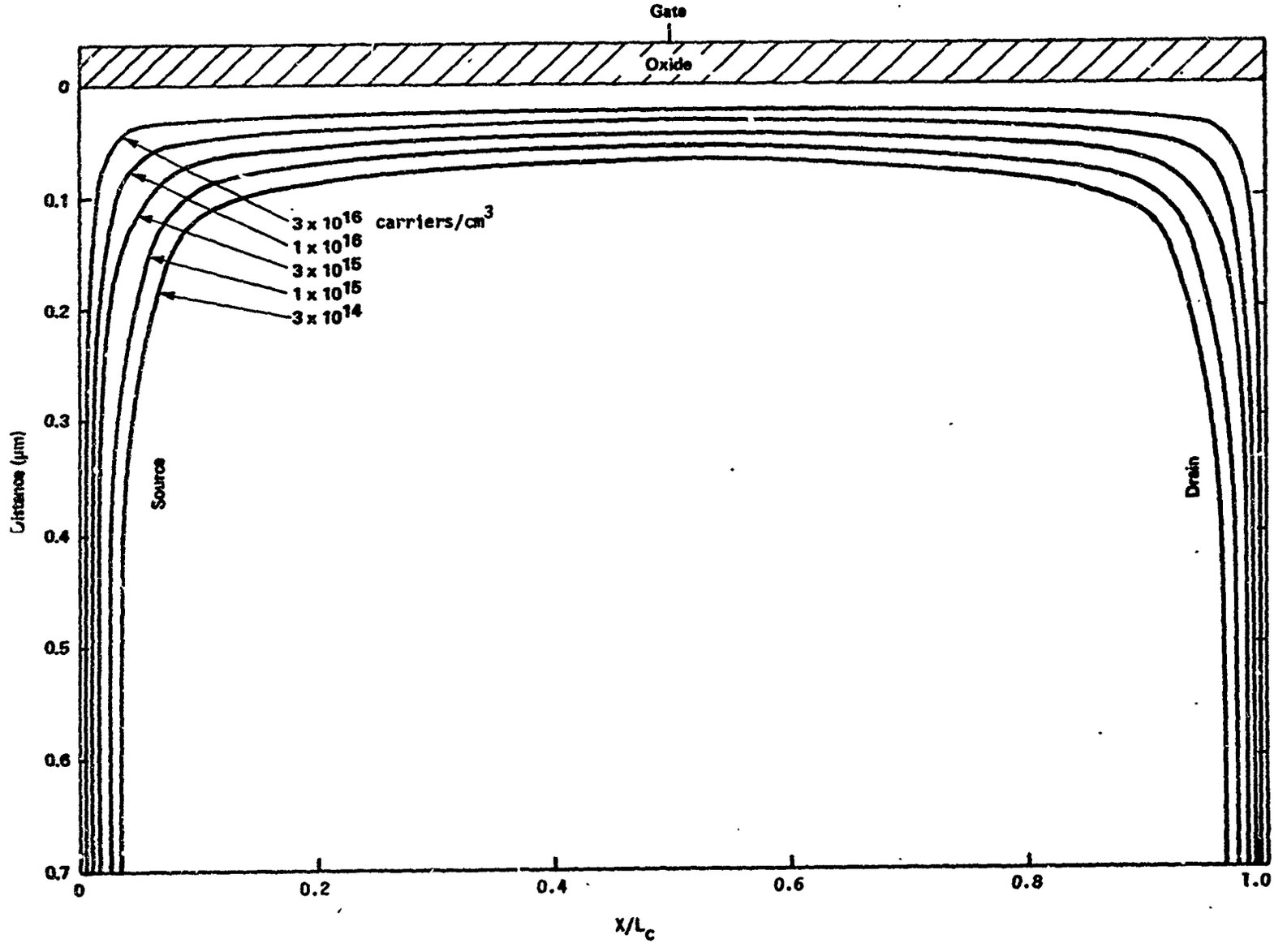

Fig. 14. Calculated minority carrier distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 1.0 volt).

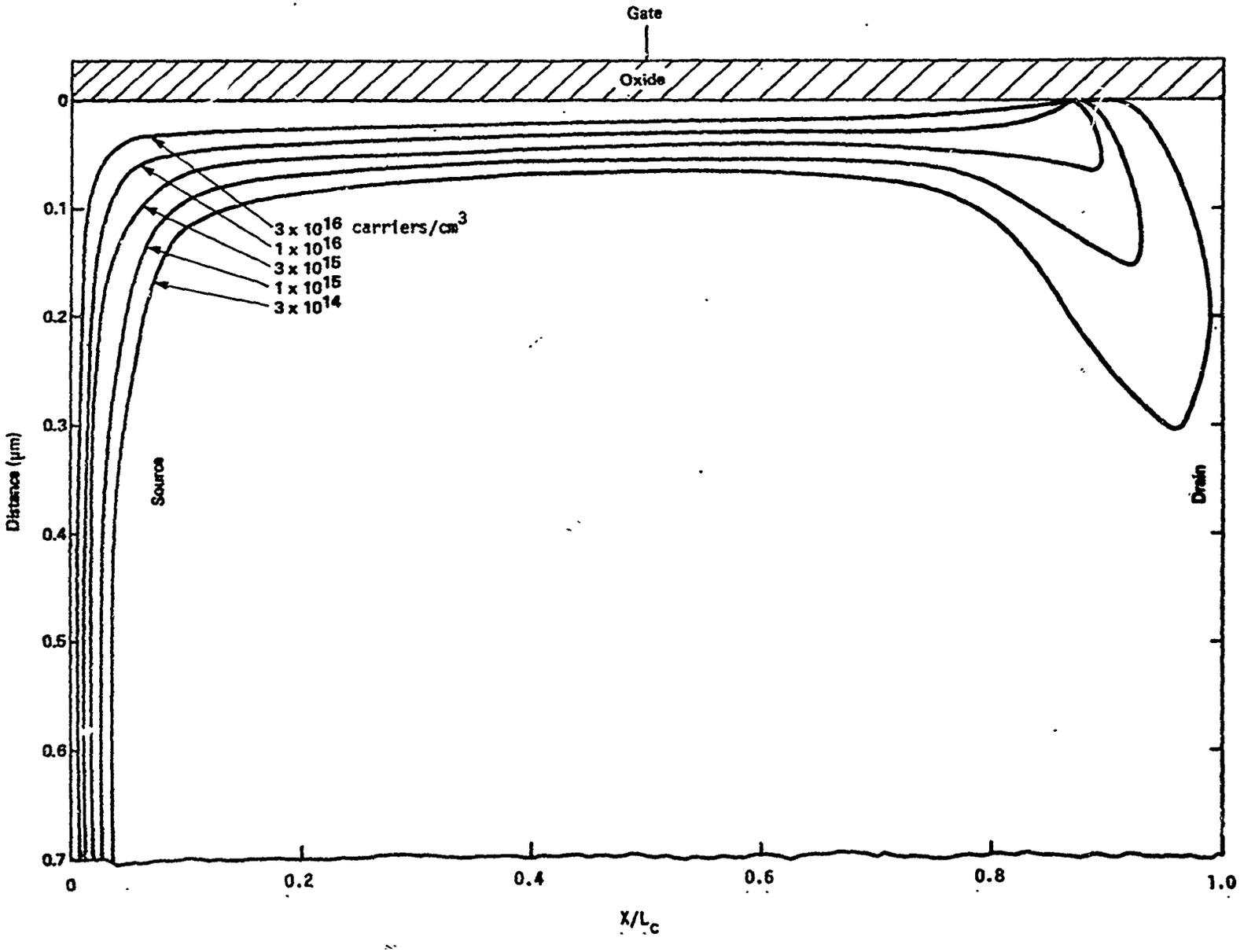

Fig. 15. Calculated minority carrier distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 6.0 volts).

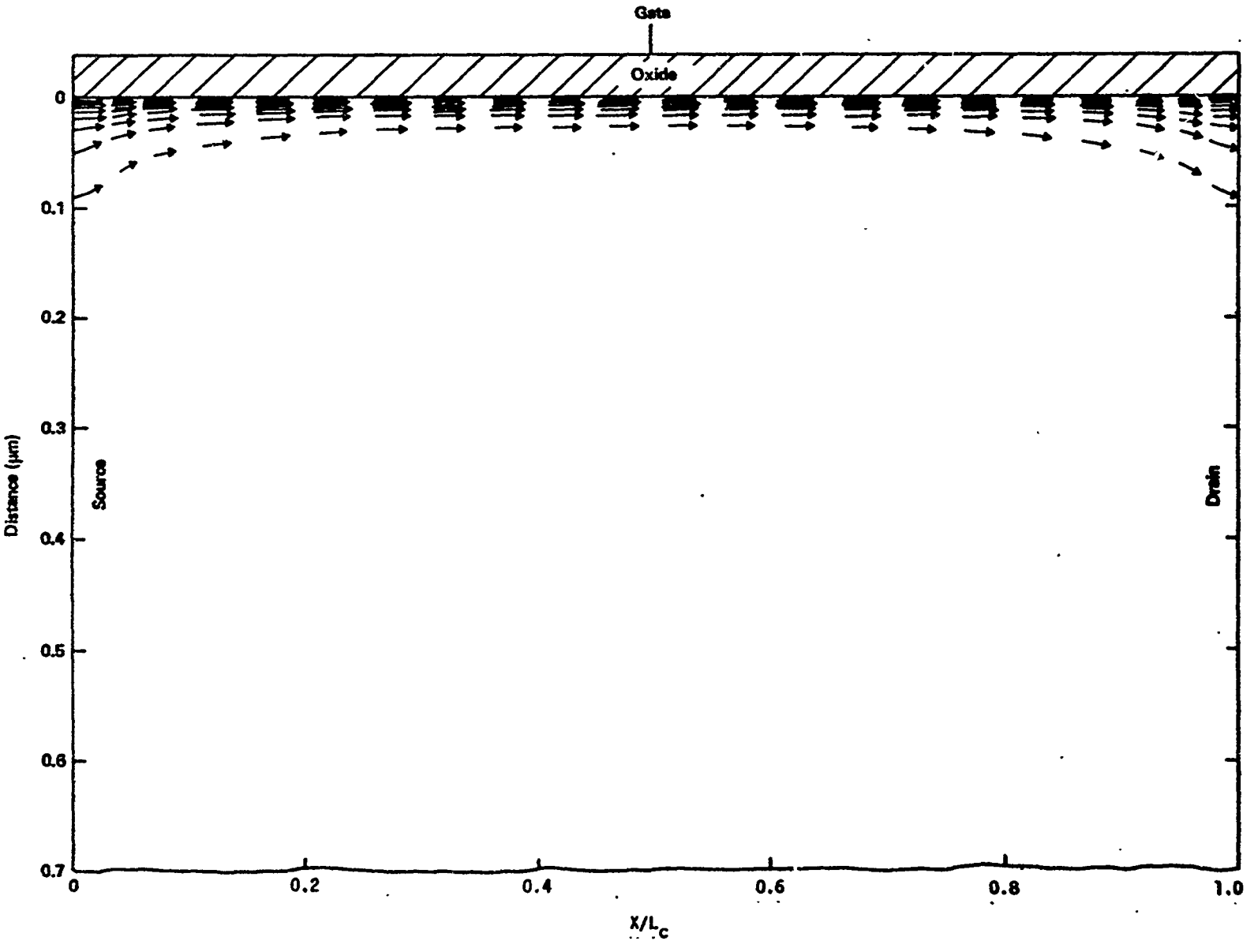

Fig. 16. Calculated mobile carrier flux distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 1.0 volt).

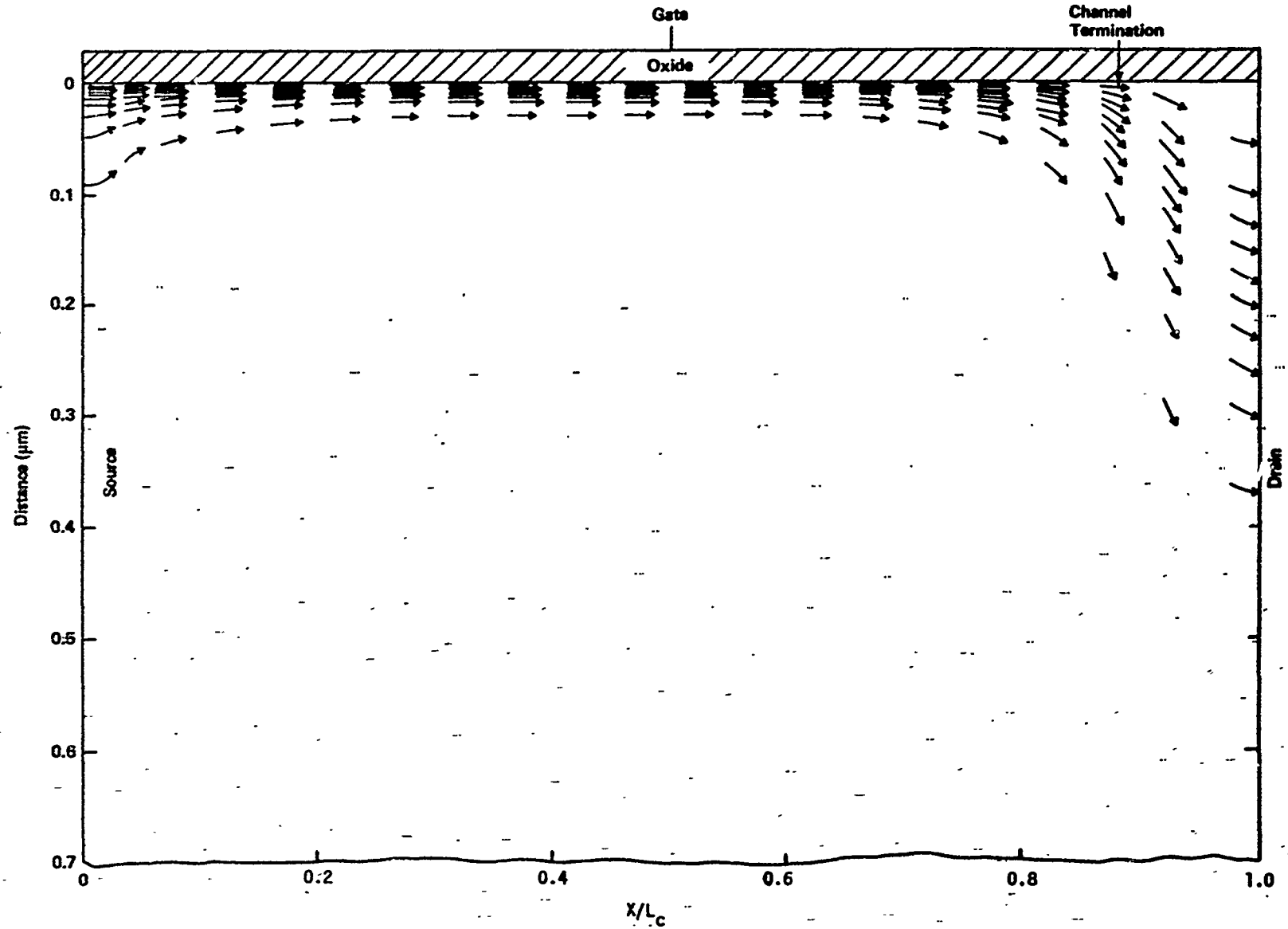

Fig. 17. Calculated mobile carrier flux distribution in an MOSFET assuming a constant carrier mobility (V_{SG}= 3.0 volts; V_{SD}= 6.0 volts).

channel pinch-off point is redesignated as a point of channel termination; thereby we are eliminating any implication that channel pinch-off produces a region of large electrical resistance, as encountered in a JFET.

2.3 Electric Current Saturation

From our calculations of IGFET operation it is concluded that two distinctly different mechanisms are capable of producing electric current saturation. In a very short channel device electric current saturation can result from velocity saturation of mobile carriers within the source-drain channel. In a long-channel IGFET velocity saturation of mobile carriers may take place, although this mechanism has little influence upon its volt-ampere characteristics. Electric current saturation in a long-channel IGFET is a consequence of the potential distribution associated with the drain junction space-charge layer. It is interesting to note, both experiment and theory show that electric current saturation can take place in an IGFET containing no velocity saturation, and when operating under circumstances that do not permit channel "pinch-off".

In a long-channel IGFET electric current saturation takes place when the applied source-drain voltage is sufficient to deplete mobile carriers from the drain end of the channel. One-dimensional theory shows this depletion can arise in two different ways: first, by a reduction of voltage between the gate electrode and the semiconductor surface and, second, by an increase of voltage between the semiconductor surface and charge neutral regions of the substrate. A reduction of gate to surface voltage produces a reduction of electric field normal to the semiconductor surface; thereby a reduction is obtained in the total surface charge within the semiconductor ($Q_i + Q_D$). Assuming the gate voltage is above threshold for the device under consideration ($V_{th} < V_{SG}$), this reduction of electric field produces an equivalent reduction of mobile charges in the inversion layer (Q_i). Similarly, one-dimensional theory also shows that if a given electric field exists at the semiconductor surface (hence, $Q_i + Q_D$ is specified) an increase of surface potential (as measured with respect to charge neutral regions of the substrate) produces an increase of depletion charge (Q_D) and, therefore, a reduction of mobile charges in the inversion layer (Q_i).

Deep saturation takes place in a long-channel IGFET when the source-drain voltage exceeds the source-gate voltage. This mode of operation produces a situation where, at some location near the drain junction, the gate induced electric field becomes zero and, hence, the source-drain channel is terminated (see Figs. 13 and 17). Between this channel termination point and the drain a negligible carrier density exists along the semiconductor surface; throughout this region the longitudinal potential distribution is determined by uncompensated ionized impurities associated with the drain junction space-charge layer. Therefore, any change of drain voltage (assuming it remains larger than the gate voltage) produces a change of voltage between the point of channel termination and the drain, and the voltage applied to the source-drain channel remains essentially constant. This mechanism was first postulated by Ihantola,[1] and has been verified by rigorous computer calculations of this semiconductor device.

Figure 18 shows the calculated volt-ampere characteristics of a typical long-channel IGFET. In this particular calculation it is assumed that mobile carriers within the source-drain channel exhibit no velocity saturation (constant drift mobility). From Fig. 18, it is evident that current saturation is initiated at a source-drain voltage that is below the value producing channel termination. This initiation of current saturation (at a reduced drain voltage) is assumed to result from the onset of channel termination, although from previous discussions it is not obvious how this situation arises in an IGFET structure.

From one-dimensional IGFET theory, the total electrostatic charge at the semiconductor surface (Q_T) is attributable to uncompensated ionized impurity atoms within a region of majority carrier depletion (Q_D) and mobile carriers within a thin inversion layer (Q_i). This total charge (Q_T) must be proportional to the normal component of electric field at the semiconductor surface,

$$\frac{\kappa_i \varepsilon_0}{W_i} (V_G - V_S) \simeq Q_i + Q_D, \qquad (2)$$

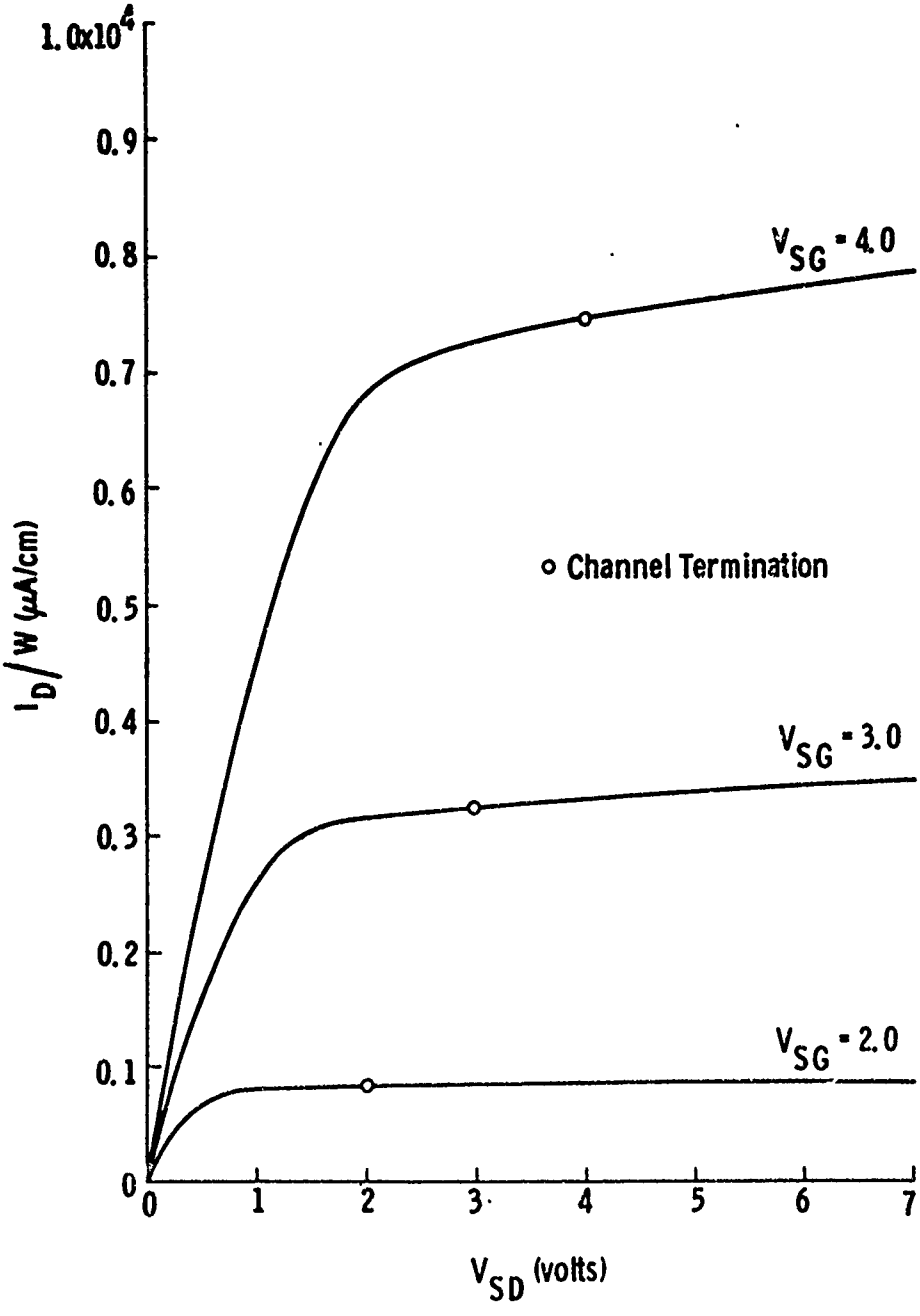

Fig. 18. Calculated volt-ampere characteristics of a long-channel IGFET assuming a constant drift mobility for carriers in the source-drain channel. (W_i=1000 Å, L_c=10 μm).

assuming the structure is free of interface charges. In addition, the total depletion charge (Q_D) is assumed to be determined by the surface to substrate voltage, giving

$$Q_D = qN_d x_D \ , \tag{3}$$

where x_D (the depletion layer width) is

$$x_D^2 = \frac{2\kappa_S \varepsilon_o V_S}{qN_D} \ . \tag{4}$$

Hence, the inversion layer charge (Q_i) is proportional to

$$Q_i \propto (V_G - V_S) \frac{\kappa_i \varepsilon_o}{W_i} - \sqrt{2q\kappa_S \varepsilon_o N_d V_S} \tag{5}$$

From eq. 5, it is implied that a variation of surface voltage (V_S), holding ($V_G - V_S$) constant, will produce a variation of inversion layer charge (Q_i). In two-dimensional IGFET theory this conclusion is not obvious. Poisson's equation in two dimensions shows that a change of voltage between the semiconductor surface and neutral substrate material does not necessarily imply a change of depletion layer charge; a redistribution of charge could arise to support this change of voltage. Specifically, in some regions ionized impurities are shared between the gate electrode and the drain (i.e. some are electrostatically tied to the gate and others to ionized impurities in the drain region). In two spatial dimensions, only those impurities electrostatically tied to the gate electrode contribute to the surface potential. For this reason, a redistribution of electrostatic association between ionized impurities in the depletion layer, the gate electrode, and the drain contact could arise with no change of the depletion charge Q_D in eq. 5.

In order to evaluate this situation the volt-ampere characteristics have been calculated for an IGFET containing an unusually large oxide thickness (Fig. 19). This large oxide thickness makes it necessary in these calculations to assume a source-gate bias in excess of about 100 volts before a significant conducting channel is formed at the semiconductor surface. From Fig. 20, at a source-gate bias of 108 volts this semiconductor device exhibits electric current saturation when the source-drain bias is in excess of about 2.0 volts; therefore (V_G-V_D) in eq. 5 is nearly constant. Although there is little (or no) experimental information on an IGFET of this type, other situations arise in the operation of more conventional devices that, to all practical purposes, experimentally verify our calculations.[17]

The calculated electrical properties of this wide gate IGFET (Fig. 20) exhibit electric current saturation, despite the fact that an essentially constant gate induced electric field exists along the semiconductor and insulator interface. From traditional theory of IGFET operation, current saturation in this structure is attributable to an increase of surface potential arising from the application of a drain voltage. An increase of surface voltage produces an increase of depletion charge (Q_D) and, hence, a decrease of inversion charge (Q_i); this decrease of inversion charge along the insulator and semiconductor interface is shown in Figs. 21 and 22. A rigorous calculation of this inversion layer carrier distribution is qualitatively consistent with one-dimensional theory, although it is necessary to verify that a decrease of Q_i is equal in magnitude to an increase of depletion layer charge (Q_D).

From a detailed study of the calculated electric field distribution near the inversion layer of this device, it has been qualitatively established that the inversion charge (Q_i) and the depletion charge (Q_D)

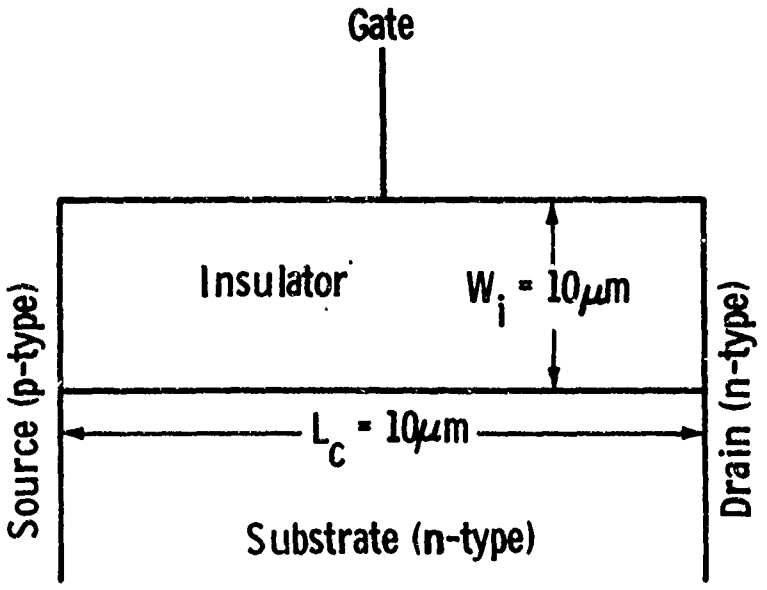

Fig. 19. Mathematical model of a wide-gate IGFET.

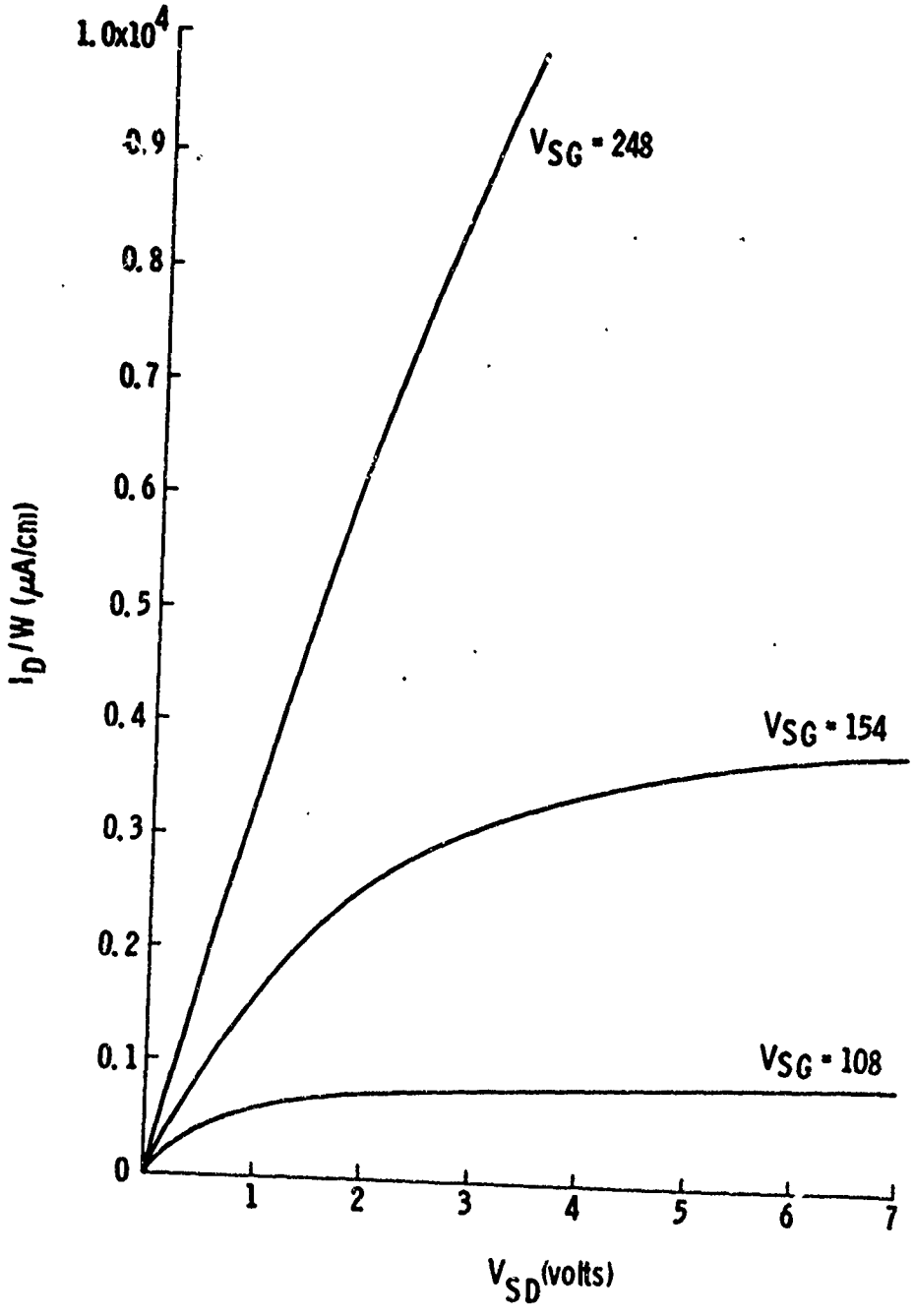

Fig. 20. Calculated volt-ampere characteristics for the wide-insulator IGFET shown in Fig. 19 (constant drift mobility).

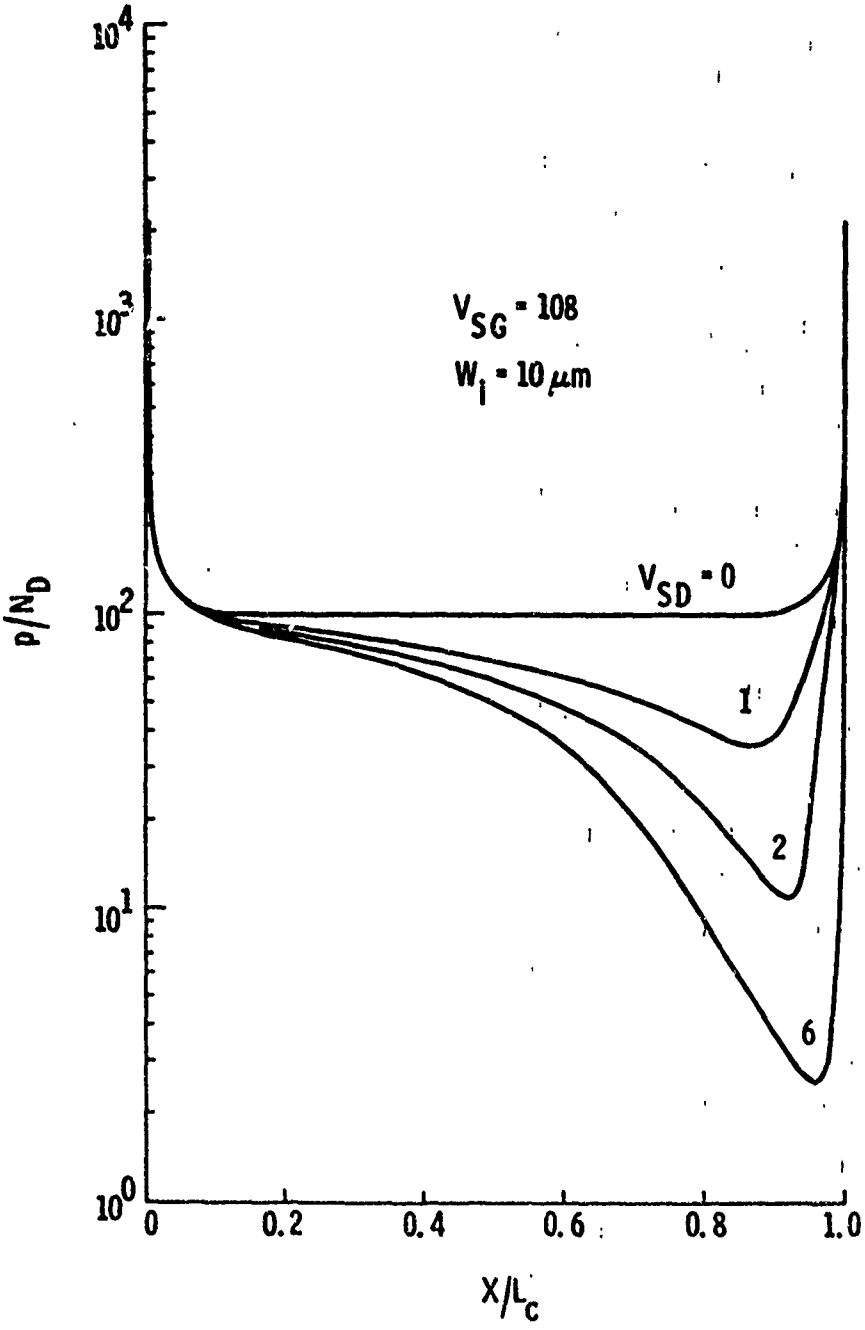

Fig. 21. Calculated mobile carrier distribution in the source-drain channel of an IGFET containing a wide insulator (Fig. 19).

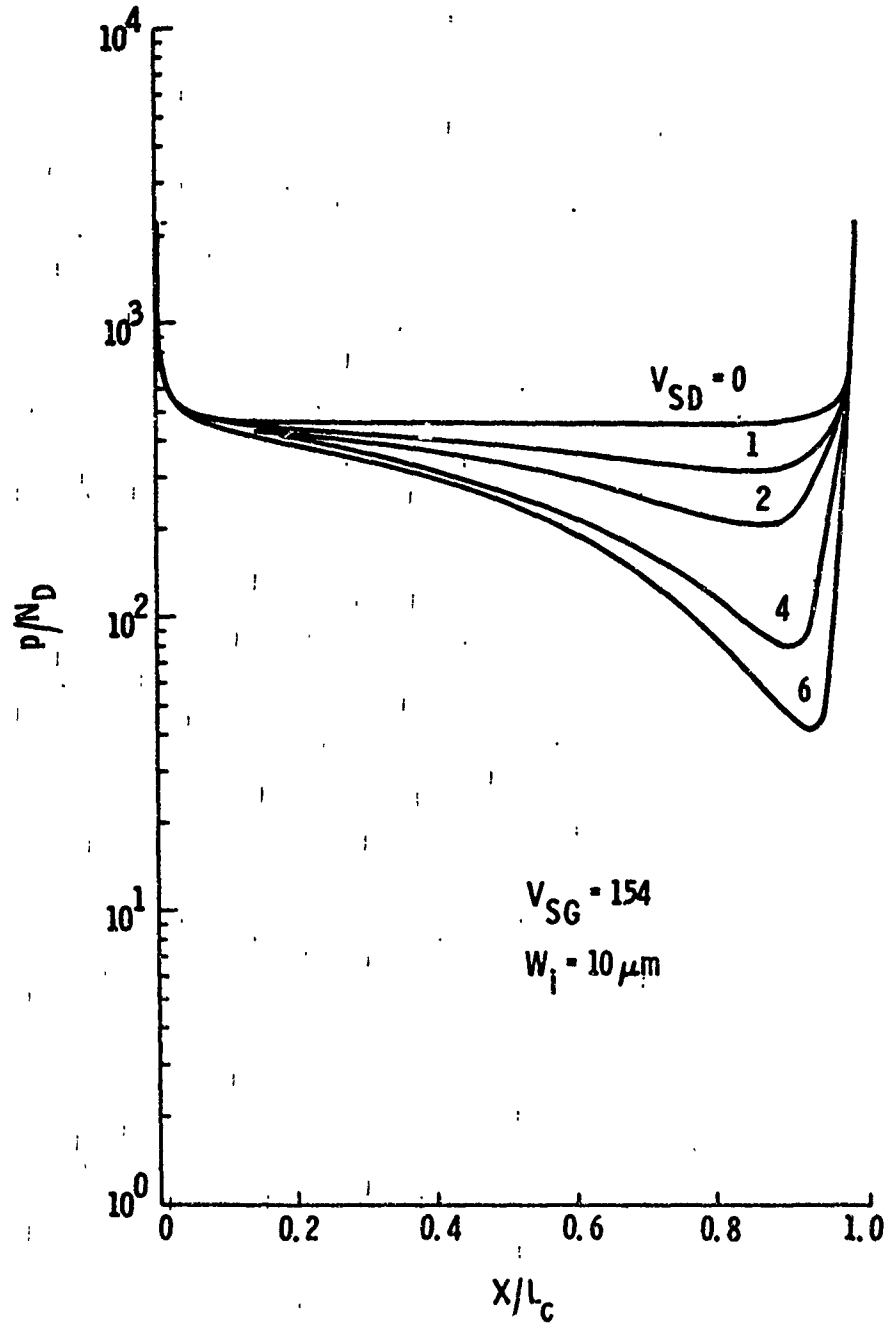

Fig. 22. Calculated mobile carrier distribution in the source-drain channel of an IGFET containing a wide insulator (Fig. 19).

satisfy the equality

$$Q_T = Q_i + Q_D \ , \qquad (6)$$

where Q_T is the total charge at the semiconductor surface. This is not intended to imply that our calculations are in agreement with traditional concepts of IGFET operation. For example, traditional theory[1,18] shows that this wide insulator structure will exhibit channel termination (or pinch-off) when the drain voltage is about 3.2 volts and the gate voltage is about 154 volts. Figure 22 shows substantial differences between our computer calculated results and the conclusions drawn from this traditional theory.

From Fig. 22, the minimum carrier density (V_{SD} = 4) is about 80 times the background doping; this magnitude is substantially larger than one would estimate from present IGFET theory. Furthermore, at a drain bias of 6.0 volts the point of minimum inversion carrier density remains close to the drain junction (Fig. 22), rather than at a location where the surface voltage is approximately 4.0 volts. This situation shows that some fundamental discrepancies exist between our traditional theory of IGFET operation and the results obtained from a rigorous solution of the problem.

Substantial effort has been directed toward the development of more accurate approximations for the mechanisms involved in electrical conduction through an inversion layer. There are many obvious simplifications in the available one-dimensional theory (gradual channel approximation, neglecting diffusion, etc.) although, thus far, by eliminating these simplifications we have only complicated the equations, with no fundamental improvement in the results. From this effort, it is concluded that important mechanisms are missing from our present one-dimensional theory of electrical conduction in a source-drain channel, yet we have thus far been unable to identify these mechanisms.

Calculations show that velocity saturation in a long-channel IGFET is restricted to the drain junction space-charge layer. Through a comparison of Figs. 23 and 24, it is concluded that velocity saturation in this region of the device has a negligible influence upon its volt-ampere characteristics. The reasons for this situation are evident from the

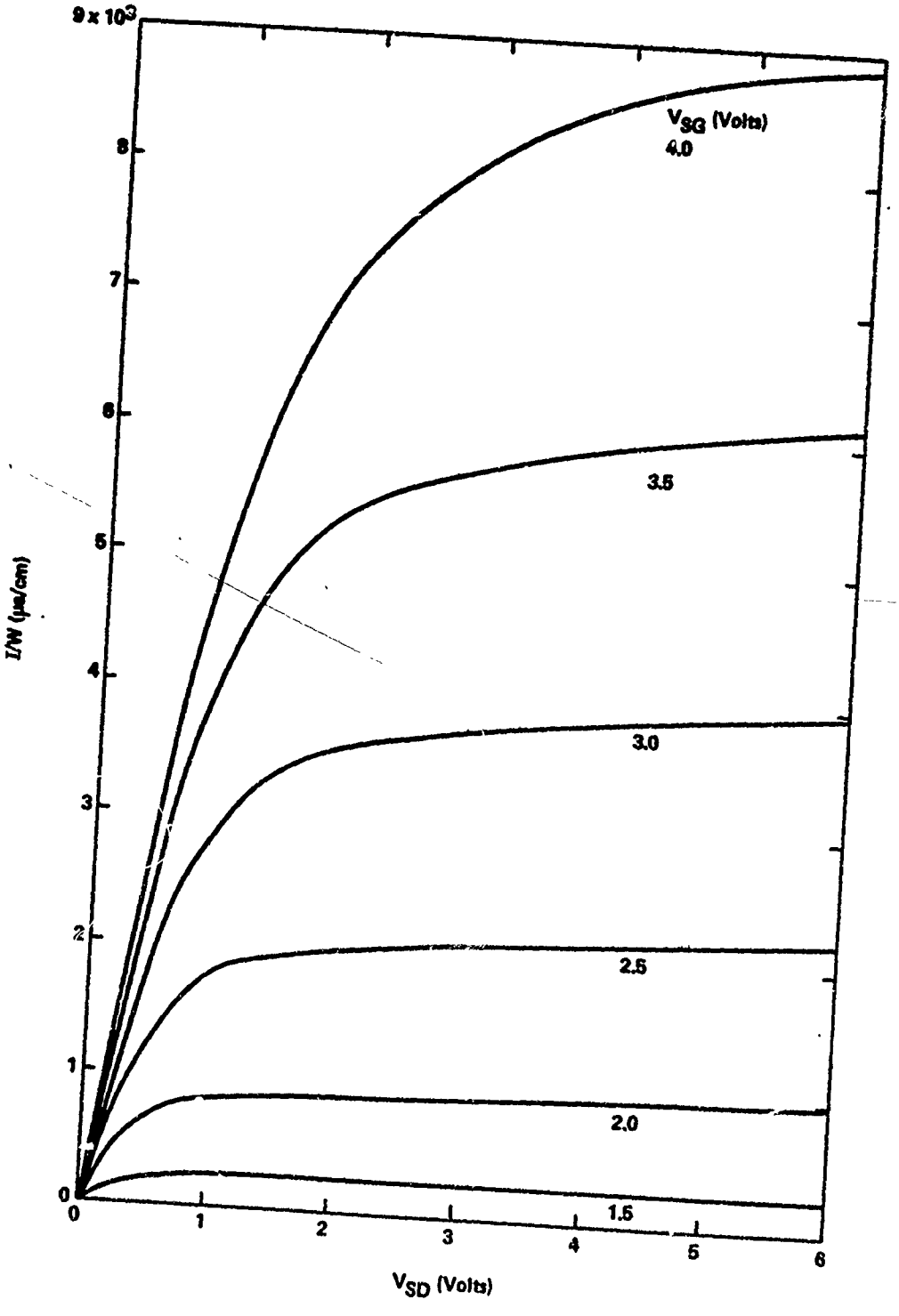

Fig. 23. Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 10 μm (constant carrier mobility).

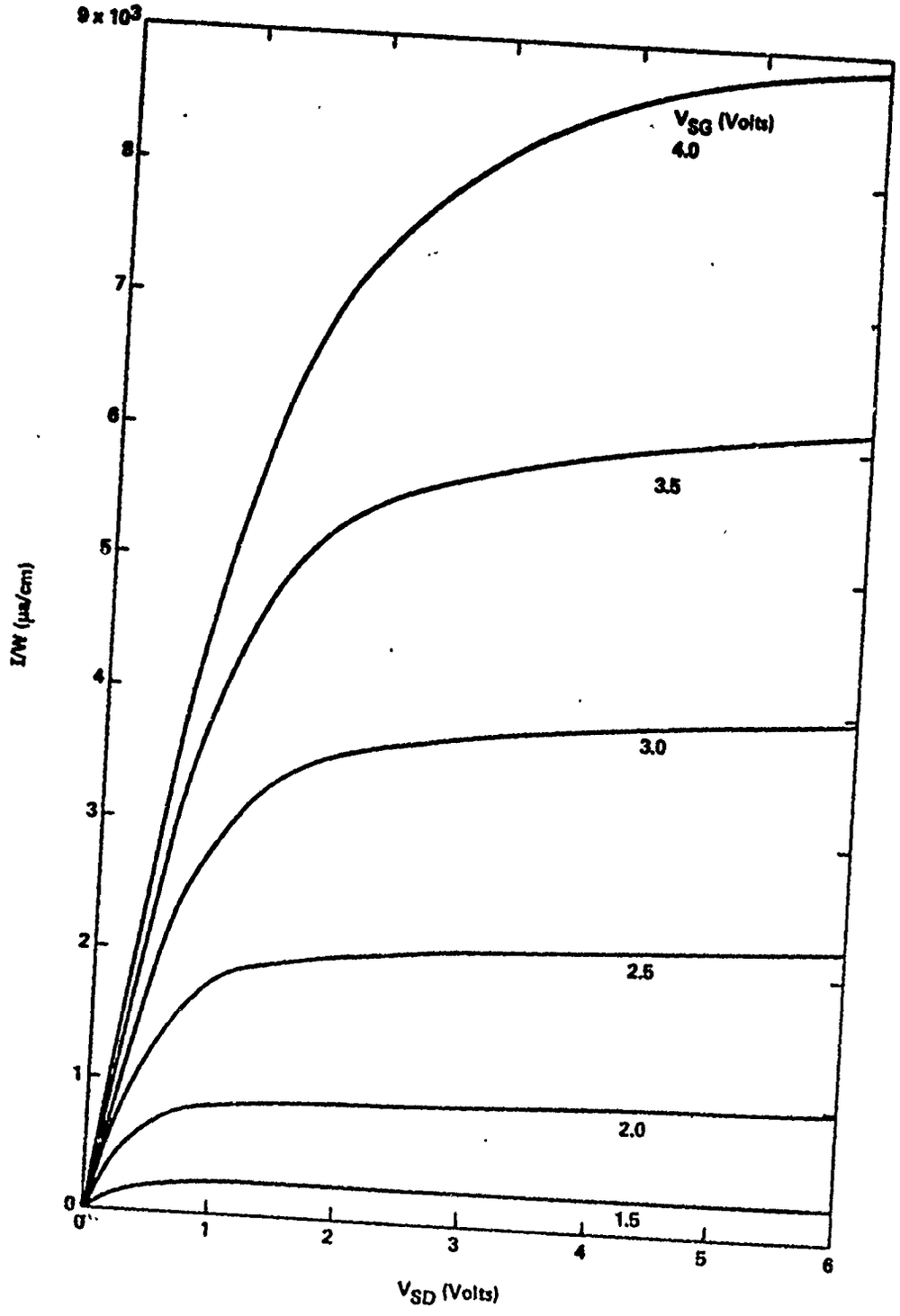

Fig. 24. Calculated volt-ampere characteristics of an MOSFET, assuming a channel length (L_c) of 10.0 μm (field-dependent carrier mobility).

calculated mobile carrier distribution, particularly in the immediate vicinity of the drain. Briefly, velocity saturation can do little to influence the electric current if, for all practical purposes, the total number of carriers producing this current can increase indefinitely. In this situation, a reduction of drift mobility (with an increase of electric field) results in an increase of carriers and, thereby, current continuity is maintained between the source and drain.

Figures 25 and 26 illustrate this mechanism, as calculated for a long-channel IGFET. A comparison between Fig. 15 (constant drift mobility) and Fig. 25 (field dependent drift mobility) shows that by introducing a field dependent mobility into these calculations a substantial increase is observed in the number of mobile carriers residing within the drain junction space-charge layer. Further, if for purposes of experiment a decrease is introduced into the maximum carrier velocity, a proportional increase is observed in the number of carriers stored within this region of the structure. As previously stated, this increase of carrier density in proportion to a decrease of carrier mobility is interpreted to represent the mechanism whereby velocity saturation has a negligible influence upon the volt-ampere characteristics of an IGFET.

This mechanism of spreading within the drain space-charge region is also observed in the calculated electron flux distribution, Fig. 26. A comparison between Fig. 17 (constant drift mobility) and Fig. 26 (field dependent drift mobility) shows a substantial spread of electron flux throughout the drain space-charge layer, with the introduction of a field dependent drift mobility.

In contrast with the long-channel structure (L_c = 10 μm), velocity saturation in a short-channel IGFET (L_c = 1.0 μm) has a significant influence upon its volt-ampere characteristics. Figure 27 shows that in the absence of velocity saturation a short channel IGFET will exhibit a relatively soft electric current saturation. A decrease of channel length, with an increase of drain voltage, is traditionally considered the mechanism producing soft current saturation, as observed in short-channel devices, although this mechanism is probably of little consequence. In an IGFET of these dimensions (L_c = 1.0 μm), velocity saturation produces a sufficient modification of the volt-ampere characteristics to probably mask any consequences of changes in channel length.

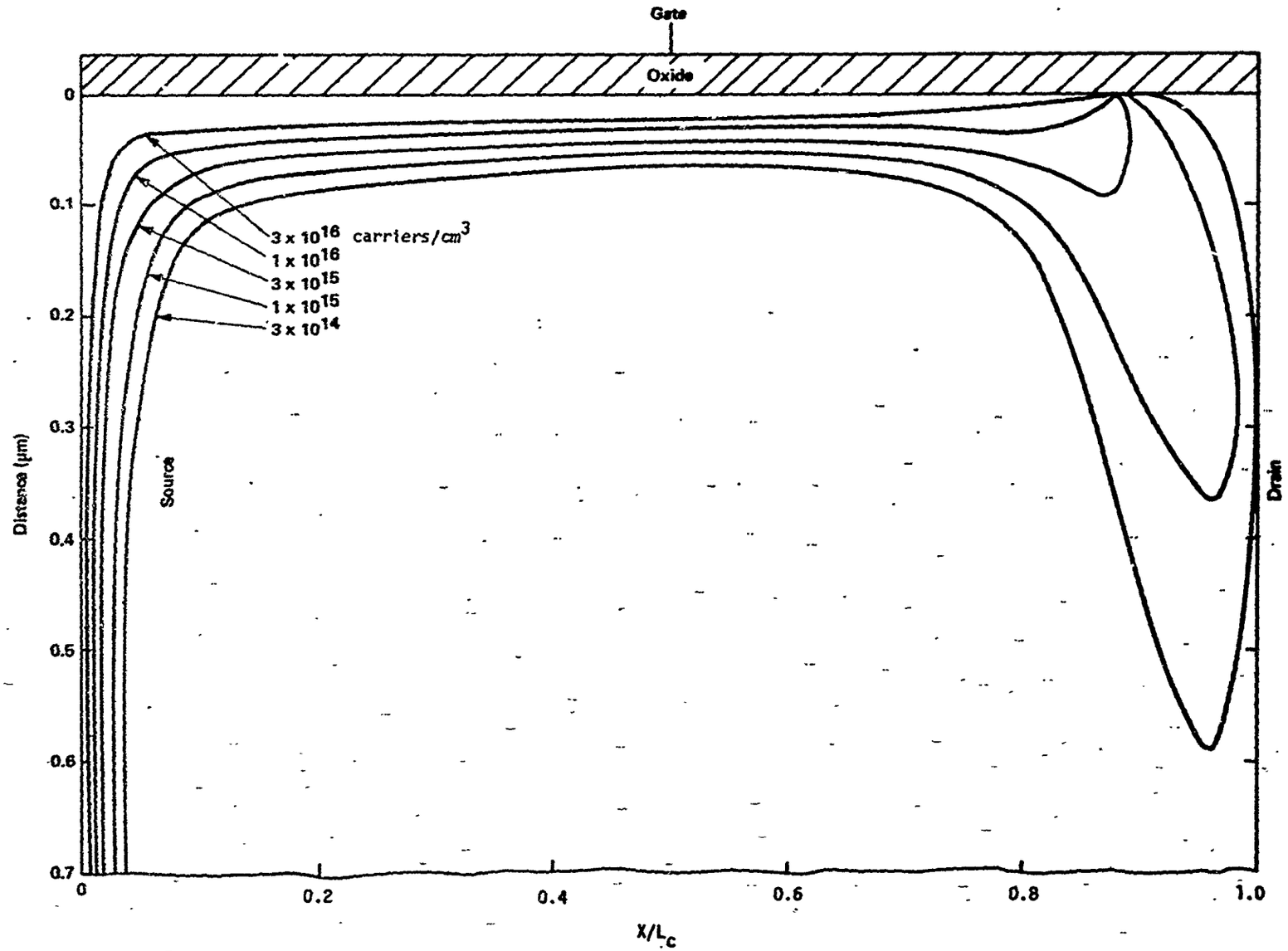

Fig. 25. Calculated minority carrier distribution in an MOSFET assuming a field dependent carrier mobility (V_{SG} = 3.0 volts; V_{SD} = 6.0 volts).

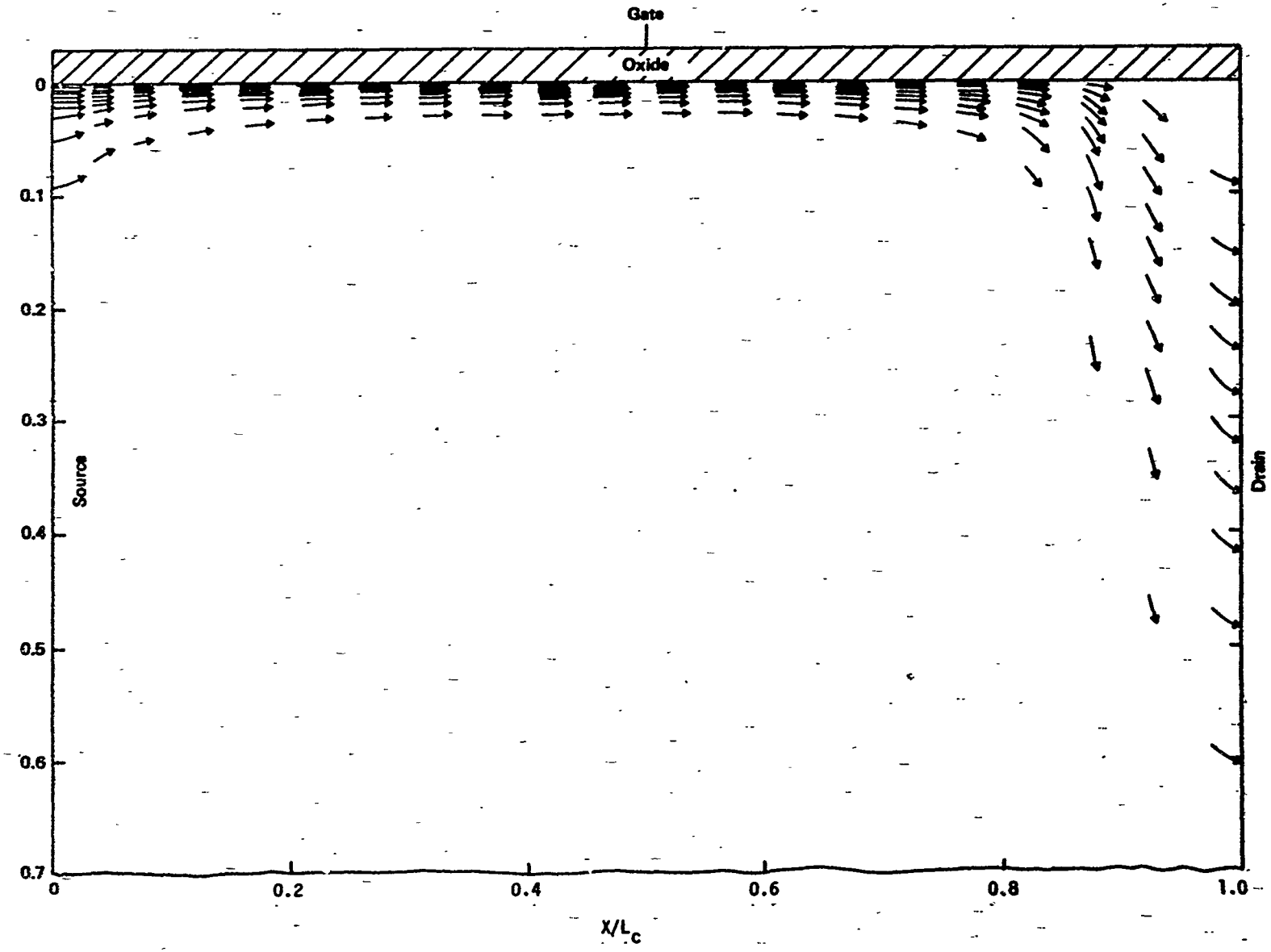

Fig. 26. Calculated mobile carrier flux distribution in an MOSFET assuming a field dependent carrier mobility (V_{SG} = 3.0 volts; V_{SD} = 6.0 volts).

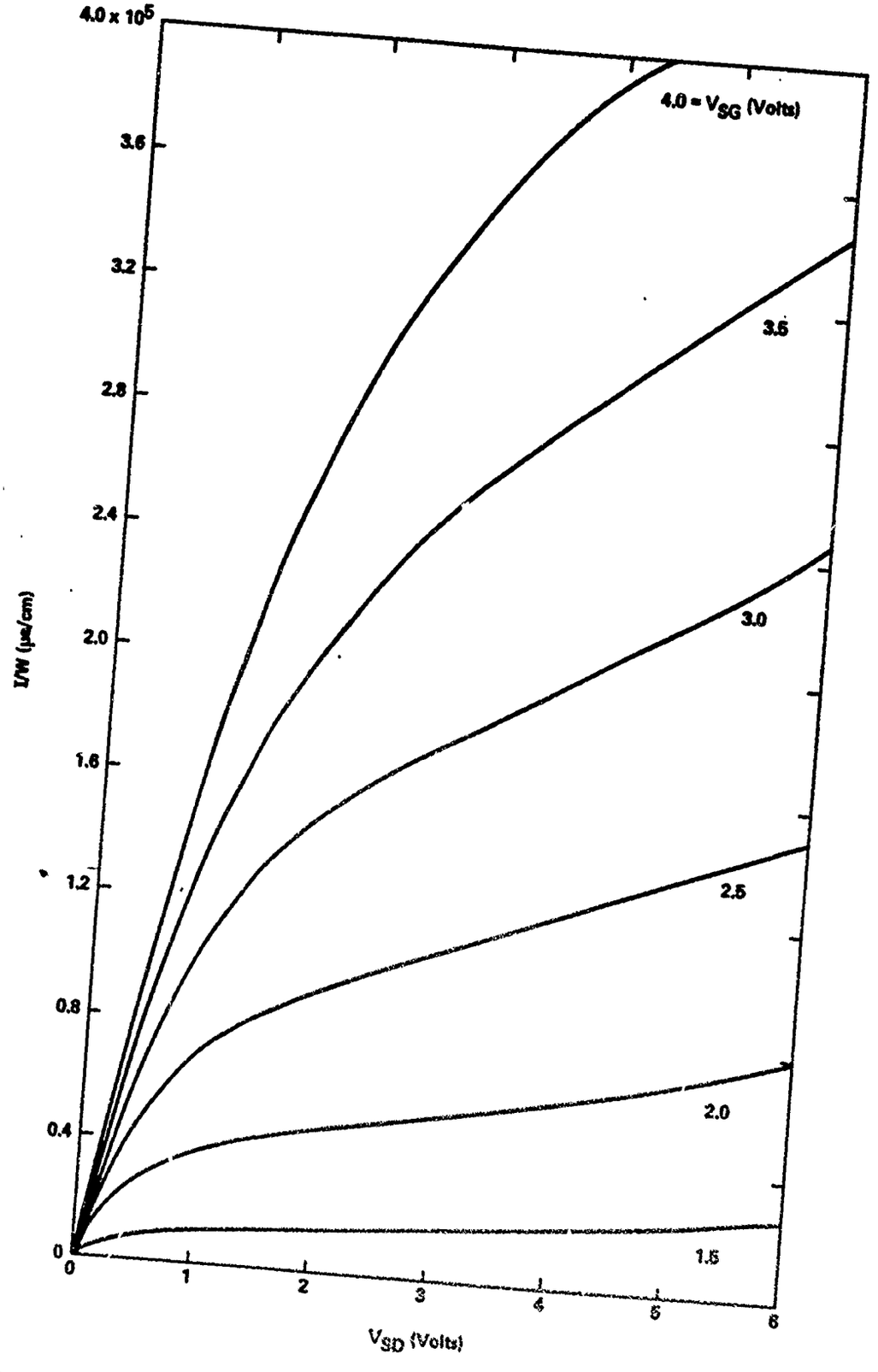

Fig. 27. Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 1.0 μm (constant carrier mobility).

Figure 28 demonstrates the influence of velocity saturation upon the volt-ampere characteristics of this short-channel device. A comparison between Fig. 27 (constant mobility) and Fig. 28 (field dependent mobility) shows that velocity saturation produces an increase of saturation resistance and, in addition, a situation where current saturation exists at a substantially lower drain voltage than can be predicted by traditional theoretical concepts; this observation is qualitatively consistent with laboratory measurements upon very short-channel IGFET structures.[19]

2.4 The Depletion Layer in an IGFET

An IGFET contains three regions of mobile carrier depletion:

1. the source junction space-charge layer,
2. the drain junction space-charge layer, and
3. a portion of the gate electrode space-charge layer.

These regions of carrier depletion have an important function. Specifically, they provide a layer of semiconductor material that electrically insulates the conducting channel from a highly conductive substrate. In addition, the electrostatic charge produced by mobile carrier depletion significantly influences the potential distribution in an IGFET. Although elementary theory treats on a one-dimensional basis all depletion layer calculations, this is clearly an approximation. The geometrical configuration of an IGFET produces regions of depletion layer interaction that must be calculated using a two-dimensional mathematical model; some initial investigations into this topic are presented here.

Figure 29 illustrates the calculated depletion layer boundary in a 5.0 μm IGFET with an assumed gate bias of 2.0 volts and an assumed drain bias of 5.0 volts. Throughout this calculation the depletion layer edge is defined as that location where the mobile carrier density is one-half the substrate impurity atom density. The overall shape of this depletion layer edge is a consequence of space-charge interaction; two depletion layers (for example the gate and drain) are located within the same region of the substrate material.

A one-dimensional mathematical model adequately characterizes the depletion layers in regions of this structute containing no

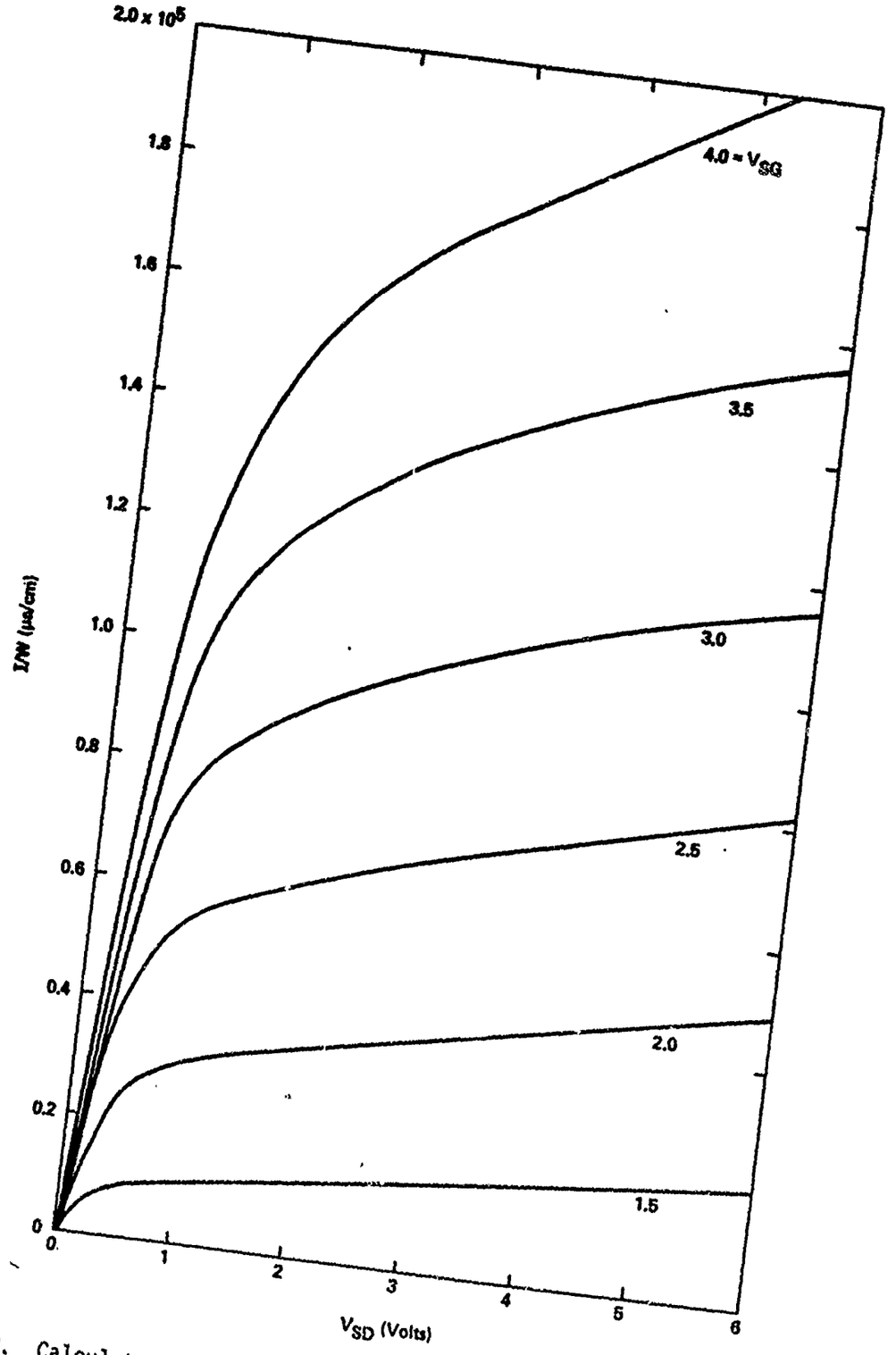

Fig. 28. Calculated volt-ampere characteristics of an MOSFET assuming a channel length (L_c) of 1.0 μm (Field dependent carrier mobility).

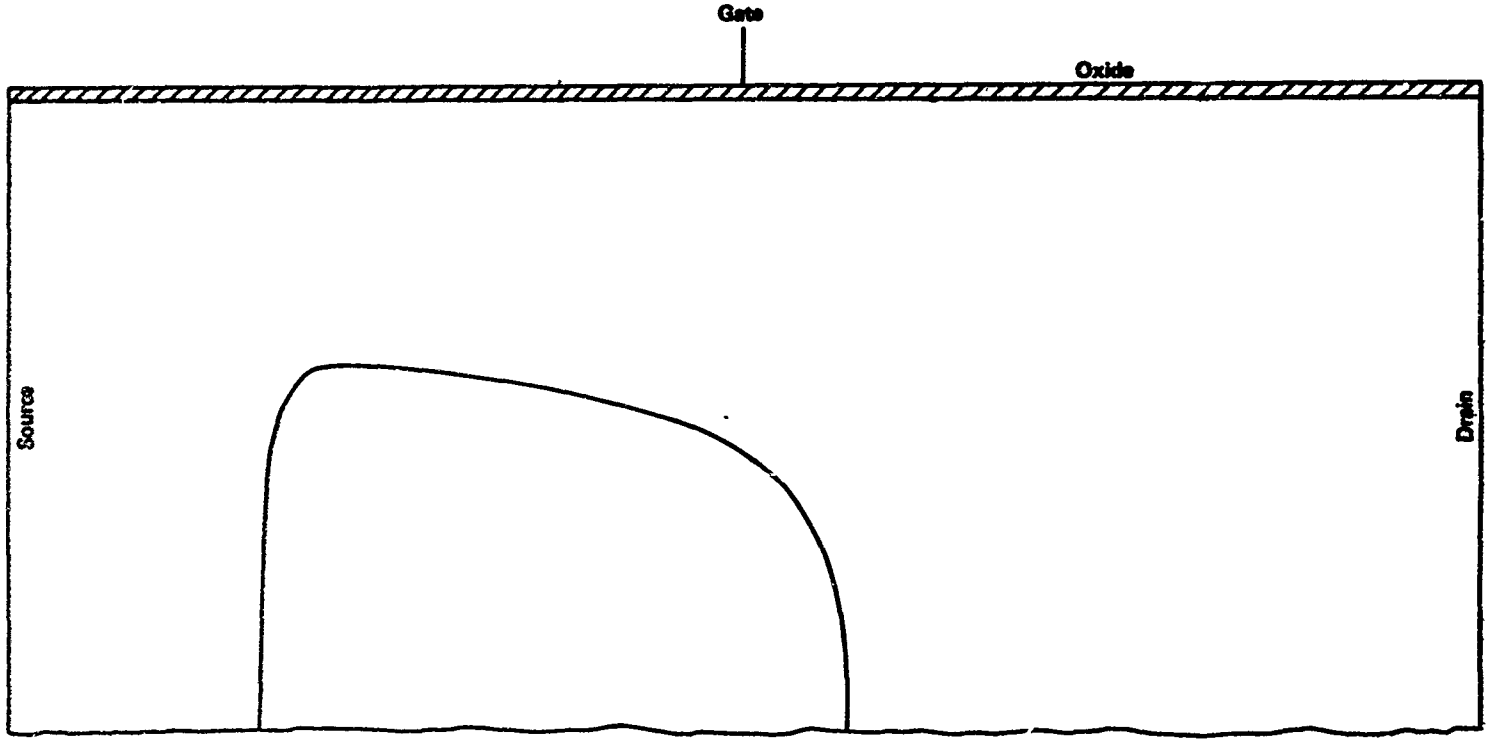

Fig. 29. Calculated depletion layer edge in a 5.0 μm MOSFET assuming V_{SG} = 4.0 volts and V_{SD} = 5.0 volts.

space-charge interaction. In contrast, space-charge layer interaction in an IGFET is a consequence of two depletion layers, at right angles to each other, sharing the ionized impurity atoms within a given region of the substrate. One consequence of this situation is a substantial reduction in the number of ionized impurity atoms available to each depletion layer; this situation is equivalent to a reduction of substrate doping. An important consequence of space-charge layer interaction is an increase of depletion layer width, and a modification of the IGFET potential distribution.

Before discussing the two-dimensional aspects of these depletion layers we shall consider the one-dimensional space-charge layer width due to the gate electrode, Fig. 30. For this calculation the source-drain voltage is zero and the depletion layer width is established for a region far removed from either the source or drain junction. As before, the space-charge layer edge is defined as that location where the majority carrier density is one-half the substrate impurity atom density.

From Fig. 30, there is no region of this curve where the depletion layer width is ideally independent of the applied gate voltage. This situation could contribute to well known inconsistencies in the definition of threshold voltage. Elementary theory of IGFET operation is based upon an assumption that the depletion layer width is constant if $V_{th} \leq V_g$; when the gate voltage exceeds threshold it is assumed all additional space-charge arises from mobile carriers within the source-drain channel. Furthermore, it is assumed that this condition of constant space-charge width is attained when the channel carrier density equals the substrate impurity atom density. For the 5.0 μm model used in this investigation, threshold should occur at a gate bias of about 0.889 V; from Fig. 30 the depletion layer width at 0.889 volts is only 90% of its maximum value. This situation could represent one source of difficulty in defining the threshold voltage.

In a two-dimensional model, the shape of interacting depletion layers can be determined from the direction and magnitude of the space-charge layer electric field. For example, regions of space-charge layer

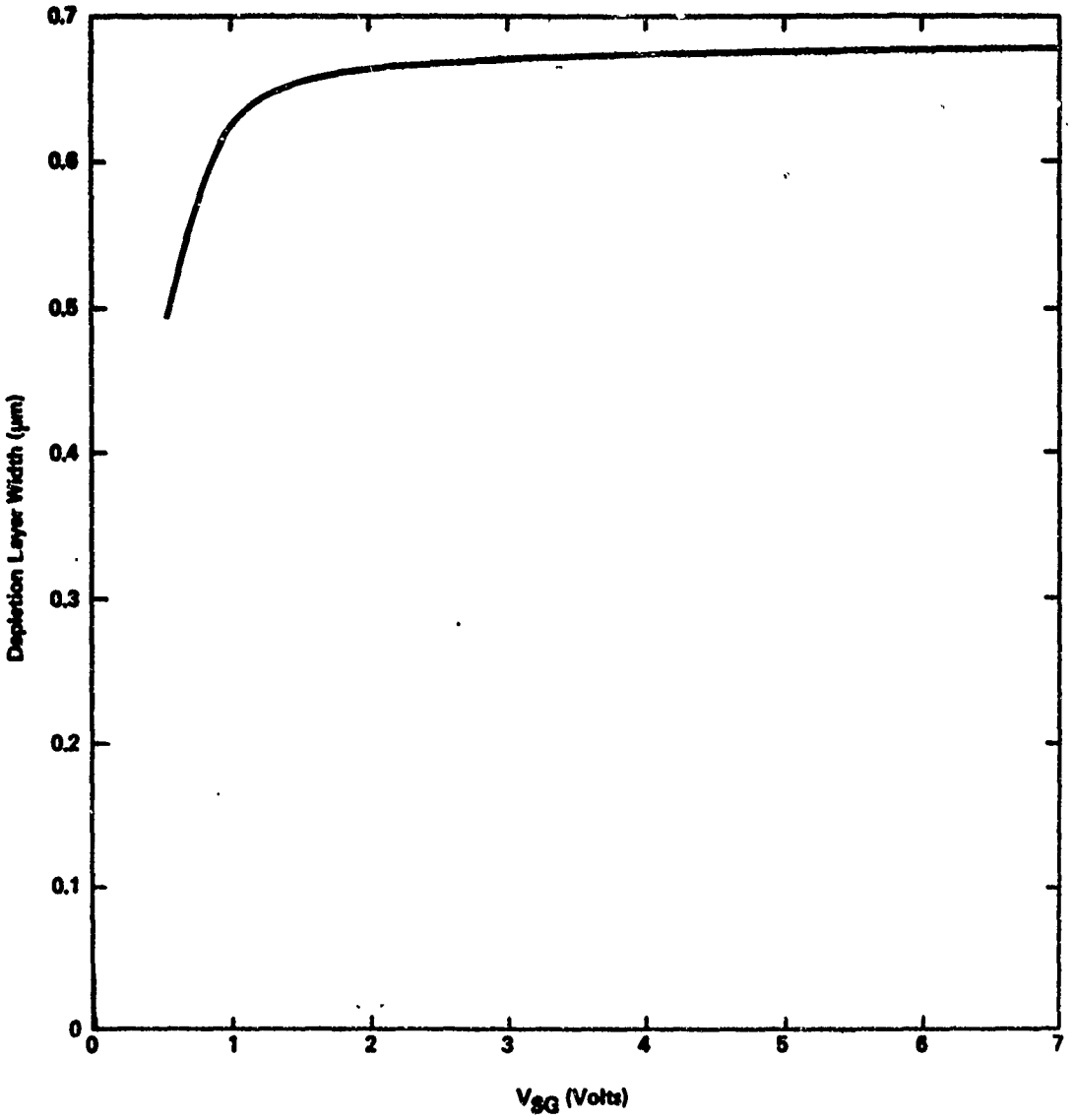

Fig. 30. Calculated gate electrode space-charge layer width in a 10.0 μm MOSFET when $V_{SD} = 0$.

interaction between the gate and drain exhibit two electric field components: one parallel and one perpendicular to the oxide and semiconductor interface. The electric field component parallel to this interface is attributable to the drain junction, and the depletion layer edge is assumed to be located where this electric field component has decreased to 1000 V/cm. Figure 31 illustrates this calculated depletion layer edge when V_{SG} = 2.0 volts and V_{SD} = 0.

Figures 32-35 show the calculated drain junction space-charge layer edge with an increase of source-drain biasing voltage. This sequence of calculations exhibit an expansion of the drain junction depletion layer into the gate electrode depletion layer. As previously stated, this expansion is attributed to the lack of ionized impurity atoms available to the drain junction; a large percentage of these ions are electrostatically tied to the gate electrode. For this reason, this gate depletion layer appears as a region of small impurity atom density through which the drain depletion layer can reach through to the source junction.

From the suggested mechanism producing reach-through, the degree of drain depletion layer expansion into the gate depletion layer should be dependent upon the source-gate voltage; this is confirmed by the calculations shown in Figs. 36-38. In this sequence of calculations (Figs. 36-38) the source-drain bias is maintained at 5.0 volts and the source-gate voltage has a magnitude of 1.0, 2.0, and 4.0 volts, respectively. At a gate bias of 1.0 volt (Fig. 36) there is little penetration of the drain space-charge layer into the gate space-charge layer. In contrast, at a source-gate bias of 4.0 volts (Fig. 38) the drain depletion layer penetrates to the source junction and produces a contraction of the source junction space-charge layer.

At this time no comparison has been made between the reach-through mechanism suggested by Figs. 32-35 and traditional source-drain punch-through. In traditional punch-through the drain space-charge layer penetrates to the source and produces a contraction of the source junction space-charge layer; this contraction is viewed as equivalent to the application of a forward bias to the source-junction. Figure 37

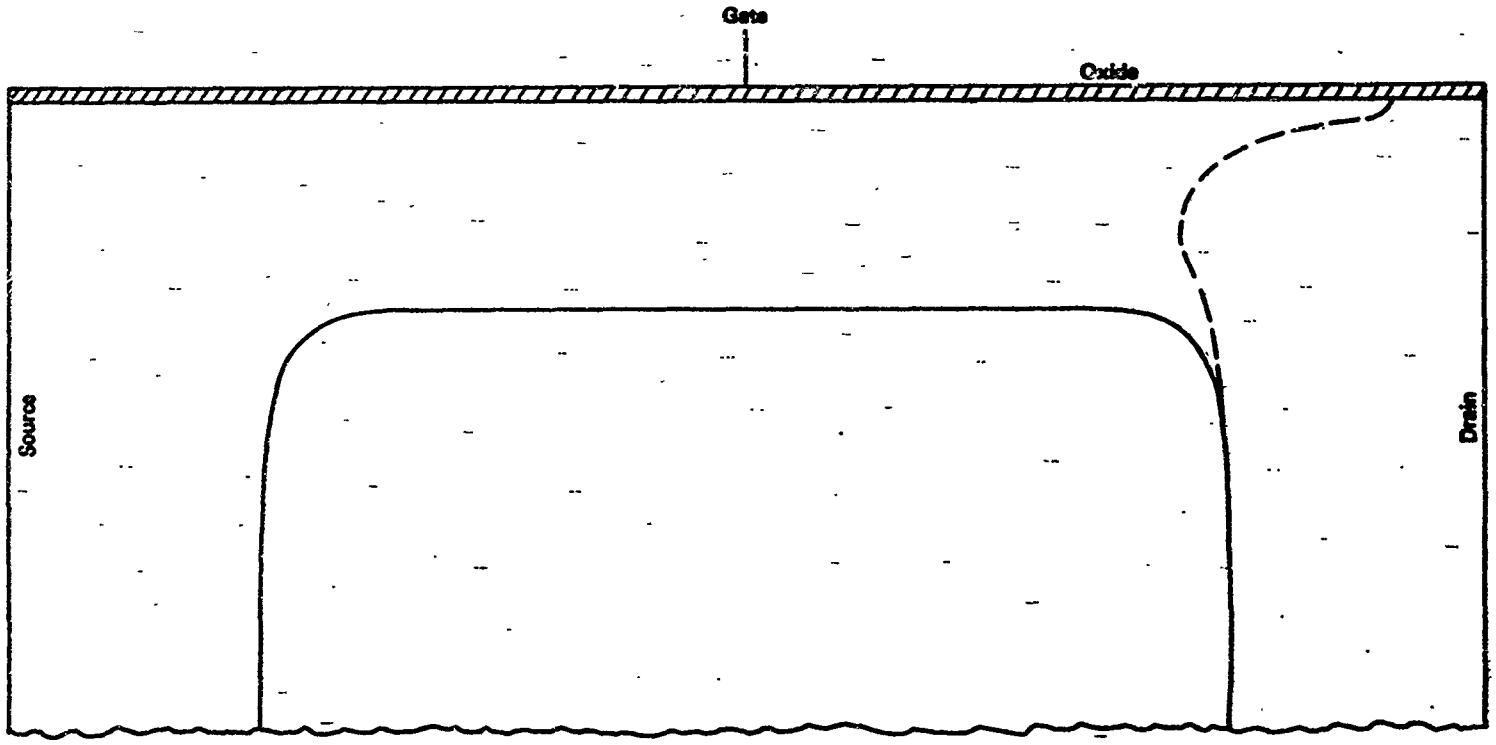

Fig. 31. Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG} = 2.0 volts and V_{SD} = 0.

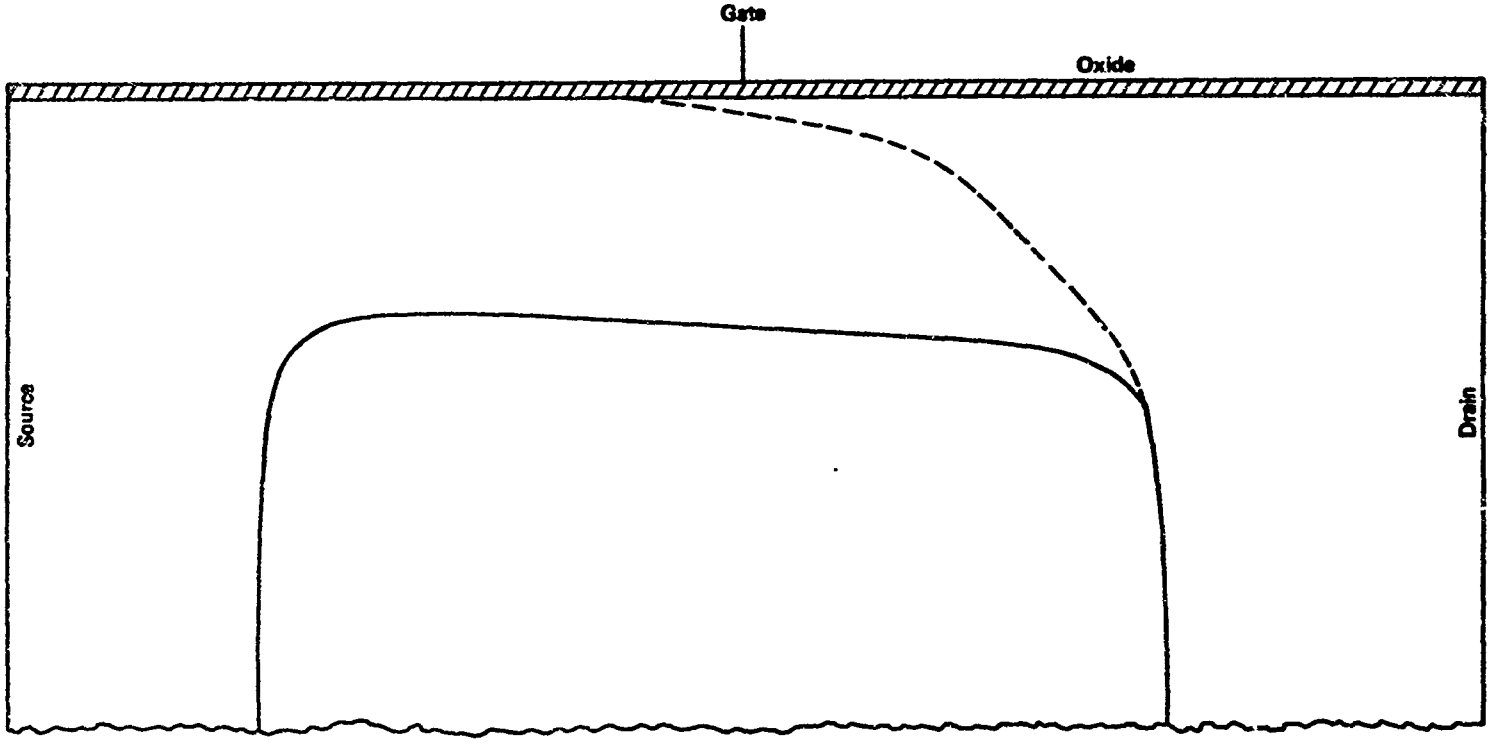

Fig. 32. Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG} = 2.0 volts and V_{SD} = 0.5 volts.

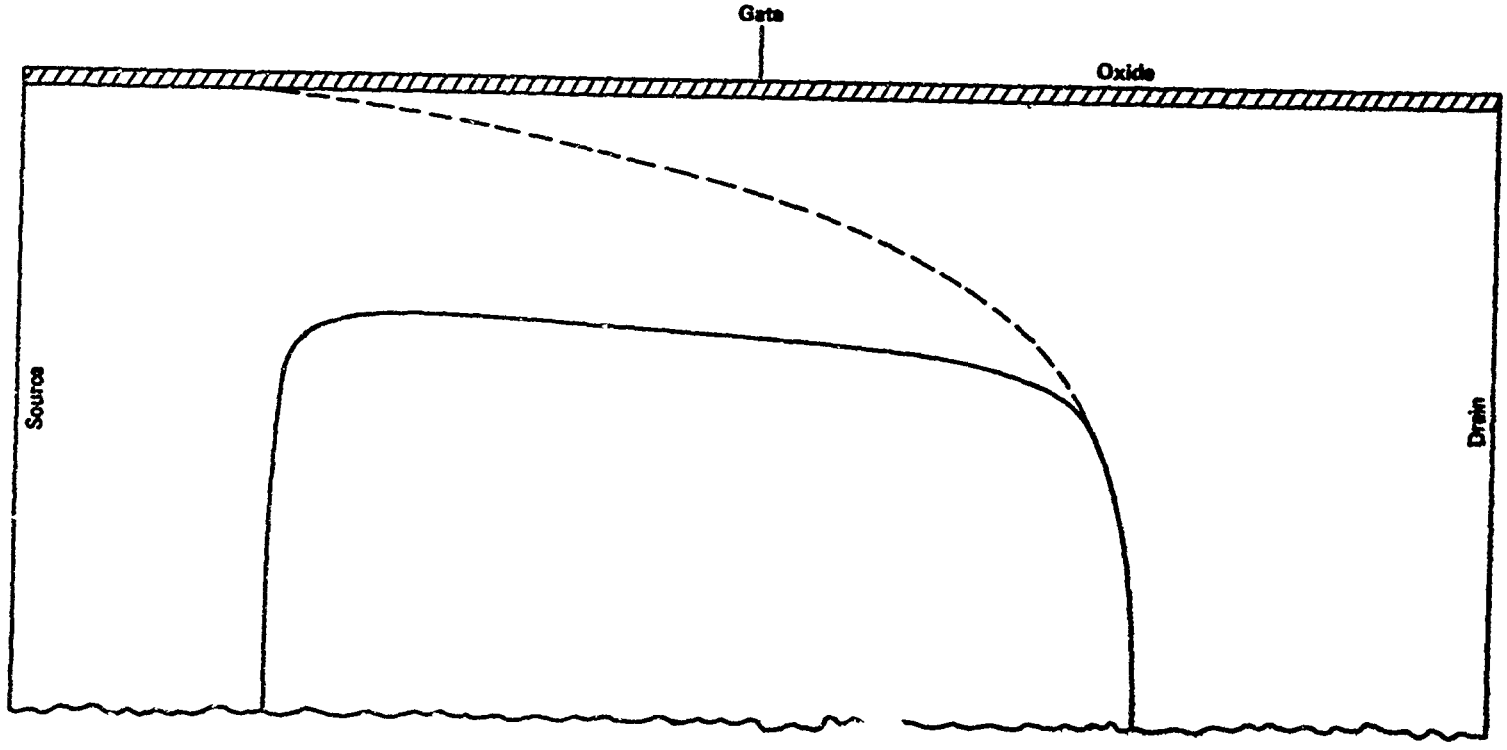

Fig. 33. Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG} = 2.0 volts and V_{SD} = 1.0 volt.

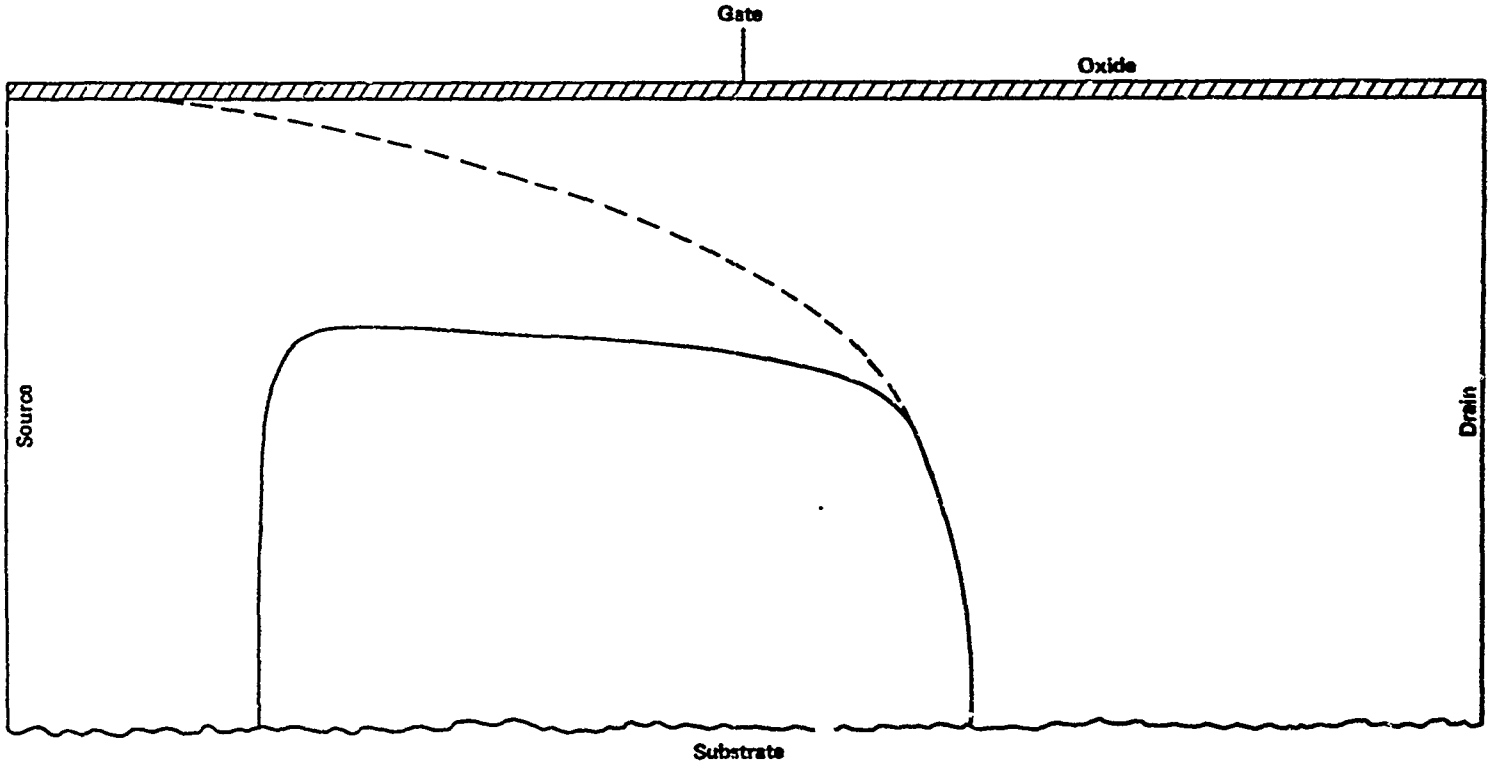

Fig. 34. Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG} = 2.0 volts and V_{SD} = 3.0 volts.

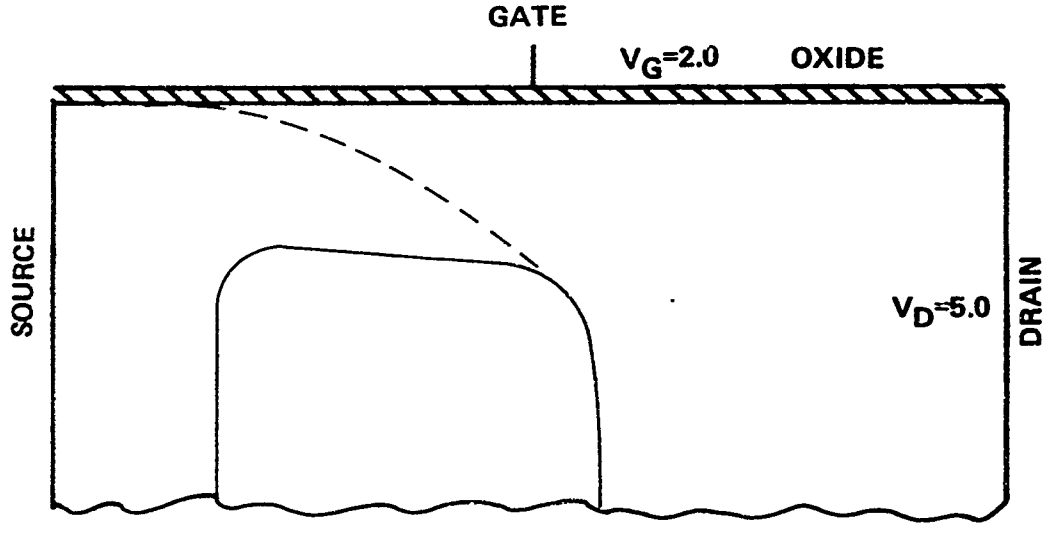

Fig. 35. Calculated drain junction depletion layer edge in a 5.0 μm MOSFET when V_{SG} = 2.0 volts and V_{SD} = 5.0 volts.

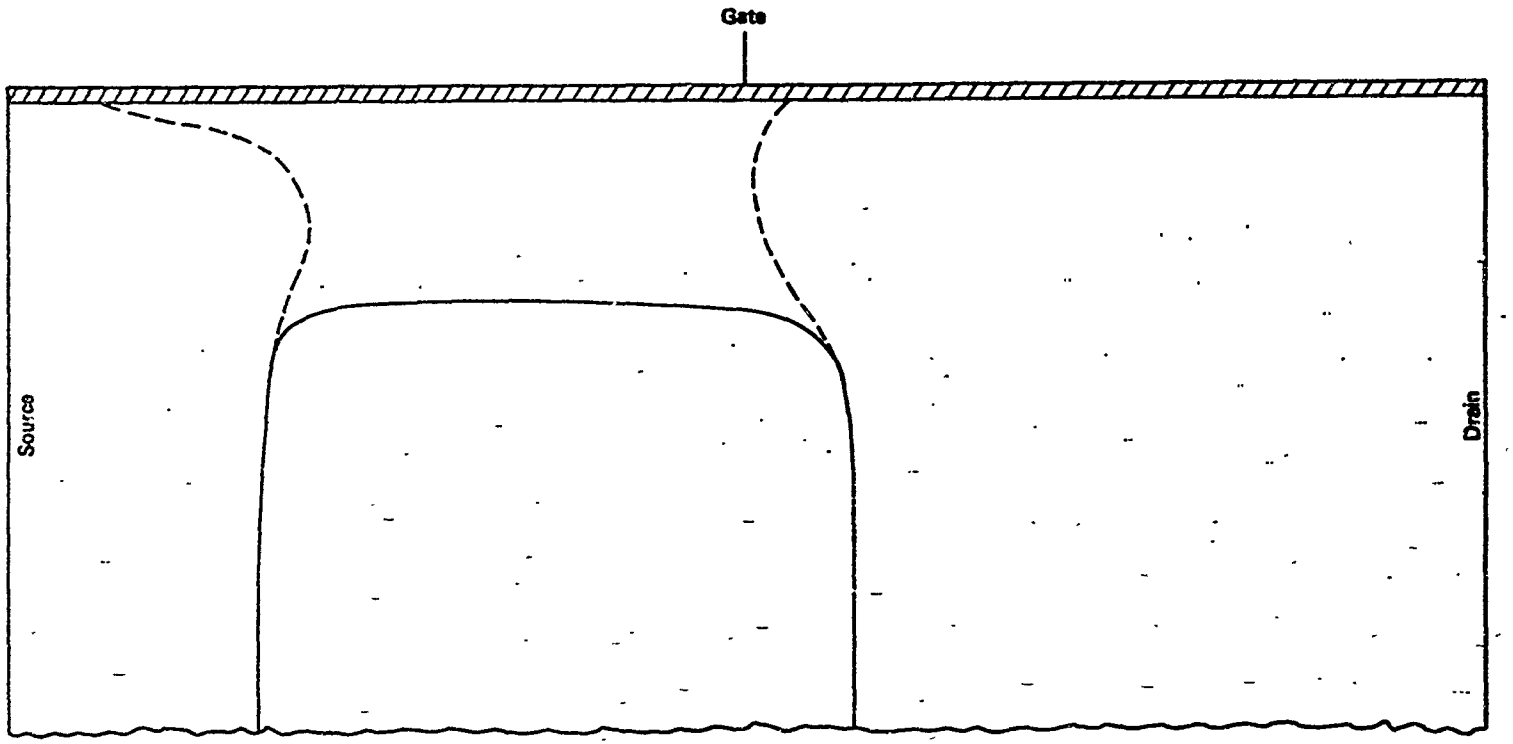

Fig. 36. Calculated source and drain depletion layer edges in a 5.0 μm MOSFET when V_{SG} = 1.0 volt and V_{SD} = 5.0 volts.

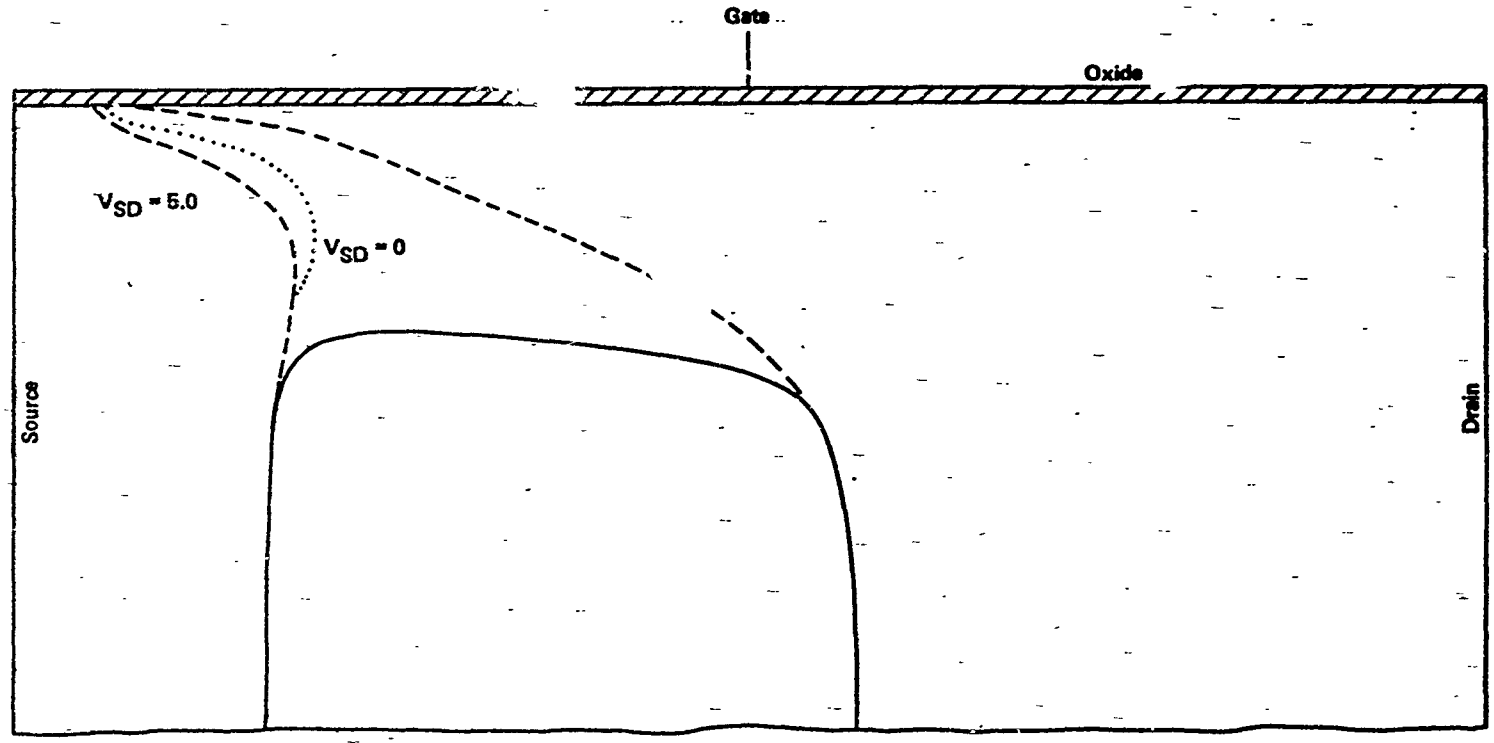

Fig. 37. Calculated source and drain depletion layer edges in a 5.0 μm MOSFET when V_{SG} = 2.0 volts and V_{SD} = 5.0 volts.

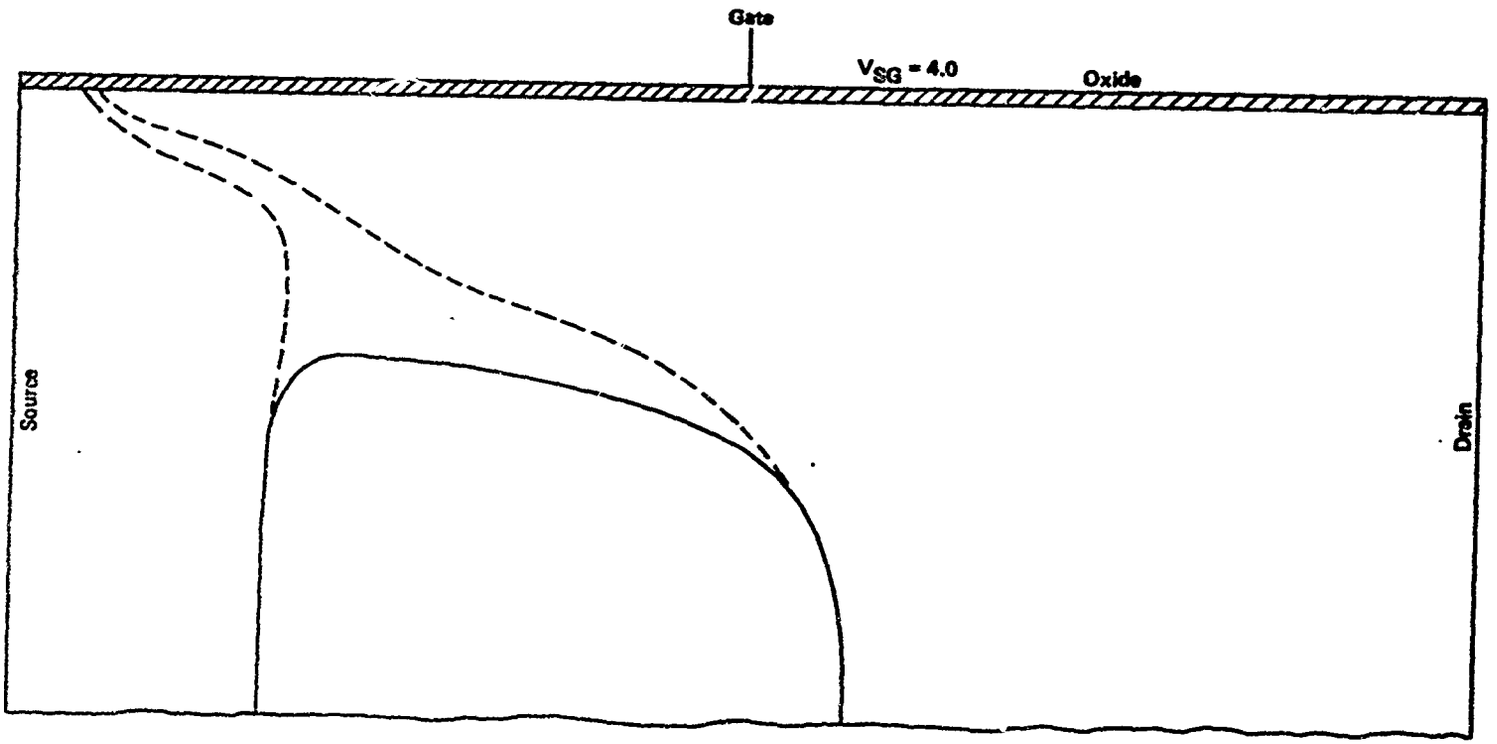

Fig. 38. Calculated source and drain depletion layer edges in a 5.0 μm MOSFET when V_{SG} = 4.0 volts and V_{SD} = 5.0 volts.

shows a comparison between the drain junction space-charge layer edge when $V_{SD} = 0$ and also when $V_{SD} = 5.0$ volts. Clearly, from Fig. 37, reach-through produces a contraction of the source junction space-charge layer, although the influence of reach-through upon the source-drain current is a topic for further investigation.

2.5 The Threshold Voltage in an IGFET

From elementary one-dimensional theory,[18] the source-drain electric current in the low current region of an IGFET volt-ampere characteristics is given by

$$I_{SD} = \eta[(V_{SG}-V_{th})V_{SD} - \frac{1}{2} V_{SD}^2] \tag{7}$$

where

$$\eta = \frac{W\epsilon_{ox}\mu}{L_c t_{ox}} \quad . \tag{8}$$

In the saturation region of these volt-ampere characteristics this source-drain electric current is given by

$$I_{SD} = I'[L_c/(L_c-L')] \tag{9}$$

where

$$L' = \left[\frac{2\epsilon_s[V_{SD}-V_{SG}+V_{th}]}{qN_D}\right]^{1/2} \tag{10}$$

and

$$I' = -\frac{\eta}{2}(V_{SG}-V_{th})^2 \quad .$$

Figure 39 presents a comparison between this elementary IGFET theory (eqs. 7 and 9) and a rigorous calculation of these volt-ampere curves. At each gate voltage the threshold voltage (V_{th} in eqs. 7 and 9) was selected to produce agreement with the computer calculated drain current at a drain biasing voltage producing deep saturation. From Fig. 39, substantial agreement is obtained between these two theoretical techniques, although to attain this agreement a different threshold voltage was required at each value of applied gate voltage.

From Fig. 39 the inferred threshold voltage changes from 1.09 volts to 1.21 volts, depending upon V_{SG}. The reasons for this situation are presently unknown; it could result from approximations used in the

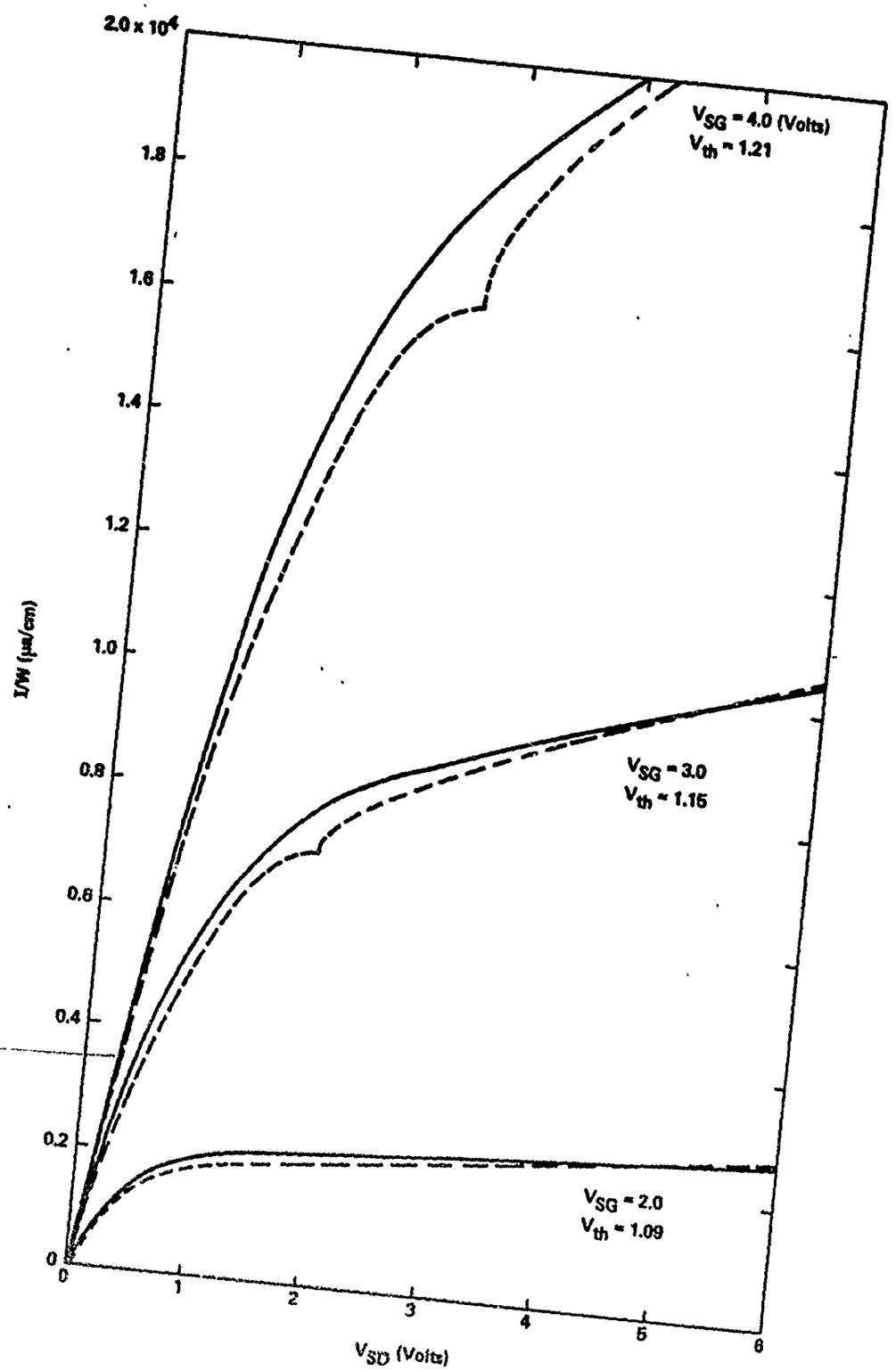

Fig. 39. Comparison between computer-calculated and volt-ampere characteristics and eqs. 7 and 9 for the 5.0 μm structure (V_{th} selected to agree in deep saturation).

development of this elementary theory or, instead, from the fact that this elementary theory remains incomplete. It appears that present methods for determining V_{th} are inconsistent and the topic requires further investigation.

2.6 The Potential Distribution in an IGFET

In the analysis of IGFET operation, the potential distribution along a source-drain channel has been subject to extensive investigation. Most workers approach this problem on a one-dimensional basis, using a gradual channel type of approximation that was initially developed by Shockley for the analysis of JFET structures. This one-dimensional analysis has been assumed adequate (without extensive verification) for structures biased well below current saturation, and for channel potential distribution calculations between the source junction and the pinch-off point (when biased into saturation). The two-dimensional aspects of this problem are recognized to be important in calculations of the potential distribution between the pinch-off point and the drain, although there is no simple method whereby this potential can be established.

The difficulties associated with solving this two-dimensional problem are evident in the publication of Chiu and Sah.[4] Specifically, Laplace's equation was solved for the region of insulating material that is approximately bounded by the channel pinch-off point and the drain. Next, this solution equation was matched to a solution of Poisson's equation for the potential distribution adjacent the insulator semiconductor boundary. In this manner, a two-dimensional analytical expression was developed for the potential distribution in the drain region of an IGFET.

It is emphasized, this mathematical method produces a very difficult problem of matching (in magnitude and derivative) two different (and complicated) solution equations at the insulator-silicon interface. A consequence of this situation is the necessity of introducing approximations that simplify this mathematical problem without decreasing the generality of the solution. In the work of Chiu and Sah, an approximation was inadvertently introduced that removed from their solution

equation the influence of ionized impurity atoms upon the calculated potential along the insulator-semiconductor boundary. The consequences of this approximation method are yet to be determined.

Figures 40 and 41 illustrate computer calculated two-dimensional potential distributions within a 5.0 μm IGFET. Figure 40 shows this potential distribution when the structure is biased below electric current saturation; in Fig. 41 the structure is assumed biased well into current saturation. The two-dimensional nature of these potential distributions is a consequence of electrostatic interactions of both the gate electrode and the drain in association with fixed lattice charges within a given region of the substrate. Regions of the drain junction far removed from the gate electrode exhibit a potential distribution that is accurately characterized by one-dimensional abrupt p-n junction theory. In contrast, near the insulator and semiconductor interface fixed lattice charges are shared between the gate and drain; therefore, the contours of constant potential are neither parallel to the gate electrode nor the drain junction.

Substantial differences can be observed between the potential distribution shown in Figs. 40 and 41. In particular, Fig. 41 shows a point along the semiconductor to oxide interface at which the semiconductor potential just equals the gate potential; on the source side of this location the electric field normal to the interface will hold mobile carriers against the semiconductor surface (thereby forming the channel) and on the drain side this field forces carriers away from the surface. It is emphasized, within both the conductive channel and the depletion layer adjacent this channel a substantial electric field can be shown to force these mobile carriers toward the drain junction.

In the source portion of the source-drain conductive channel (Fig. 41) the electric field at the silicon-oxide interface forces mobile carriers against the semiconductor surface; this field is attributed to an electrostatic interaction between the gate junction and electrostatic charges within both the channel (mobile carriers) and the depletion layer (ionized impurity atoms). In the drain portion of the source-drain channel this electric field is in the opposite

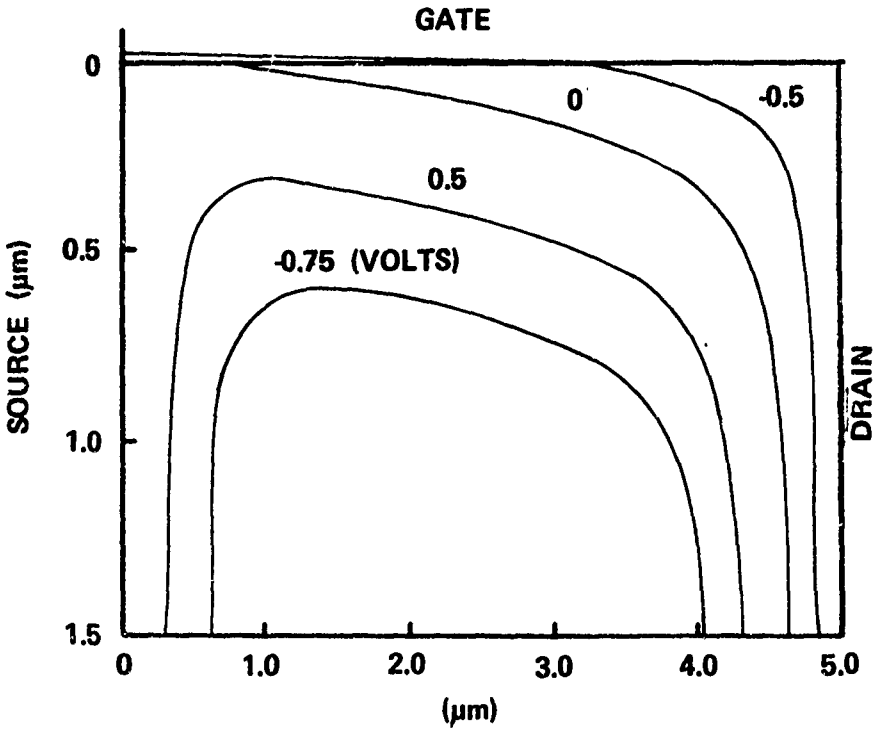

Fig. 40. Calculated (eq. 1) potential distribution in an MOSFET $N_D = 2 \times 10^{15}$ atoms/cm^2; $V_{SD} = 1.0$ volt; $V_{SG} = 3.0$ volts.

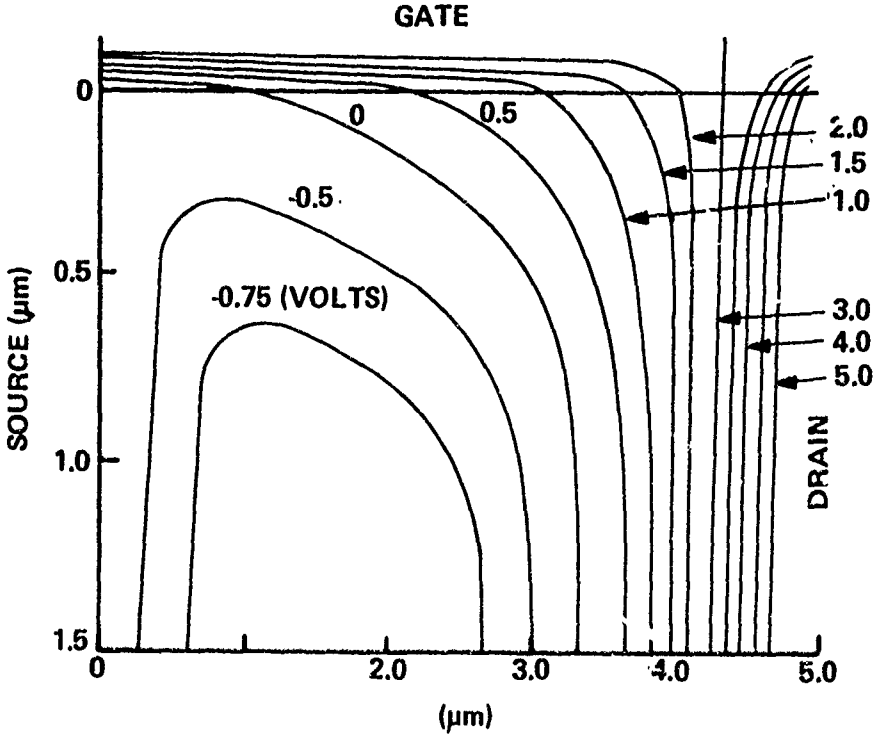

Fig. 41. Calculated (eq. 1) potential distribution in an MOSFET $N_D = 2 \times 10^{15}$ atoms/cm^2; $V_{SD} = 6.0$ volts; $V_{SG} = 3.0$ volts.

direction; it forces mobile carriers away from the semiconductor surface. Clearly, this reversal of electric field results from an electrostatic polarity reversal between the gate electrode and the semiconductor material. As a consequence, in this drain region the electrostatic charges seen by the gate electrode must be of opposite polarity to those seen in the source region of this channel. Figure 42 shows how this is accomplished.

Figure 42 illustrates the calculated contours of constant relative electrostatic displacement between the gate electrode and drain. The electric field attributable to this mechanism, in conjunction with the electric field produced by ionized impurities in the semiconductor material, produce the potential distribution shown in Fig. 41 (at the drain end of the channel). From Fig. 42 some charges within the drain region are electrostatically tied to the gate electrode, and thereby we obtain the required reversal of electric field.

From Figs. 40 and 41, it becomes evident that several weaknesses should be observed in a one-dimensional approximation for the potential distribution along a source-drain channel. For example, at large values of drain voltage this approximate theory becomes questionable: at voltages sufficient to produce substantial drain junction space-charge layer expansion into the gate depletion layer. Similarly, between the point of channel termination and the drain junction one-dimensional theory takes into consideration only the potential distribution arising from the drain junction space-charge layer. These weaknesses of the approximate theory become evident in Figs. 43-48.

Figures 43-45 present a comparison between one-dimensional and two-dimensional calculations of source-drain channel potential distribution for the 10.0, 5.0 and 1.0 μm structure, respectively. From these illustrations, substantially better agreement is obtained at small values of source-drain voltage. It is suspected that this agreement could be significantly improved if the source junction space-charge layer potential distribution was included in the one-dimensional approximation. At large values of drain voltage the one-dimensional theory is shown to be incomplete in its characterization of the channel potential

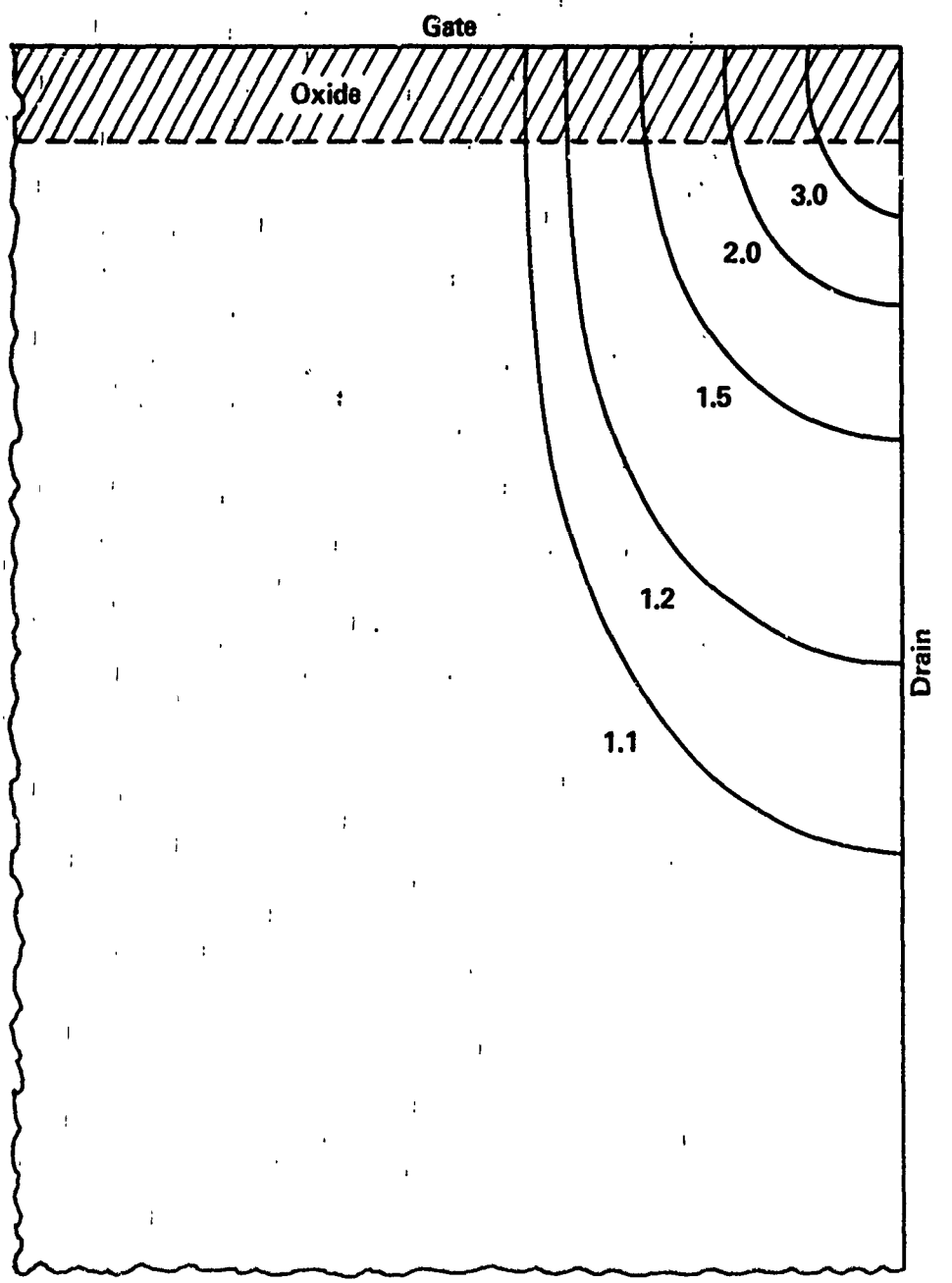

Fig. 42. Calculated contours of constant relative electric displacement between the gate electrode and the drain.

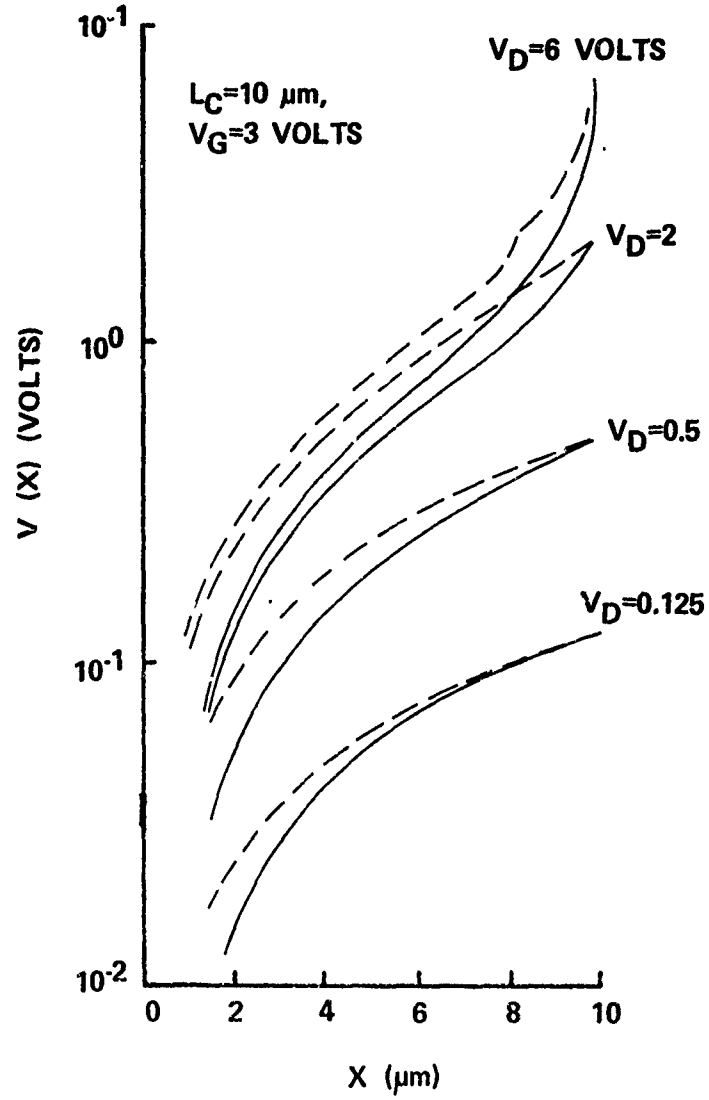

Fig. 43. Calculated potential distribution along the oxide-silicon interface in a 10.0 μm MOSFET (V_{SG} = 3.0 volts).

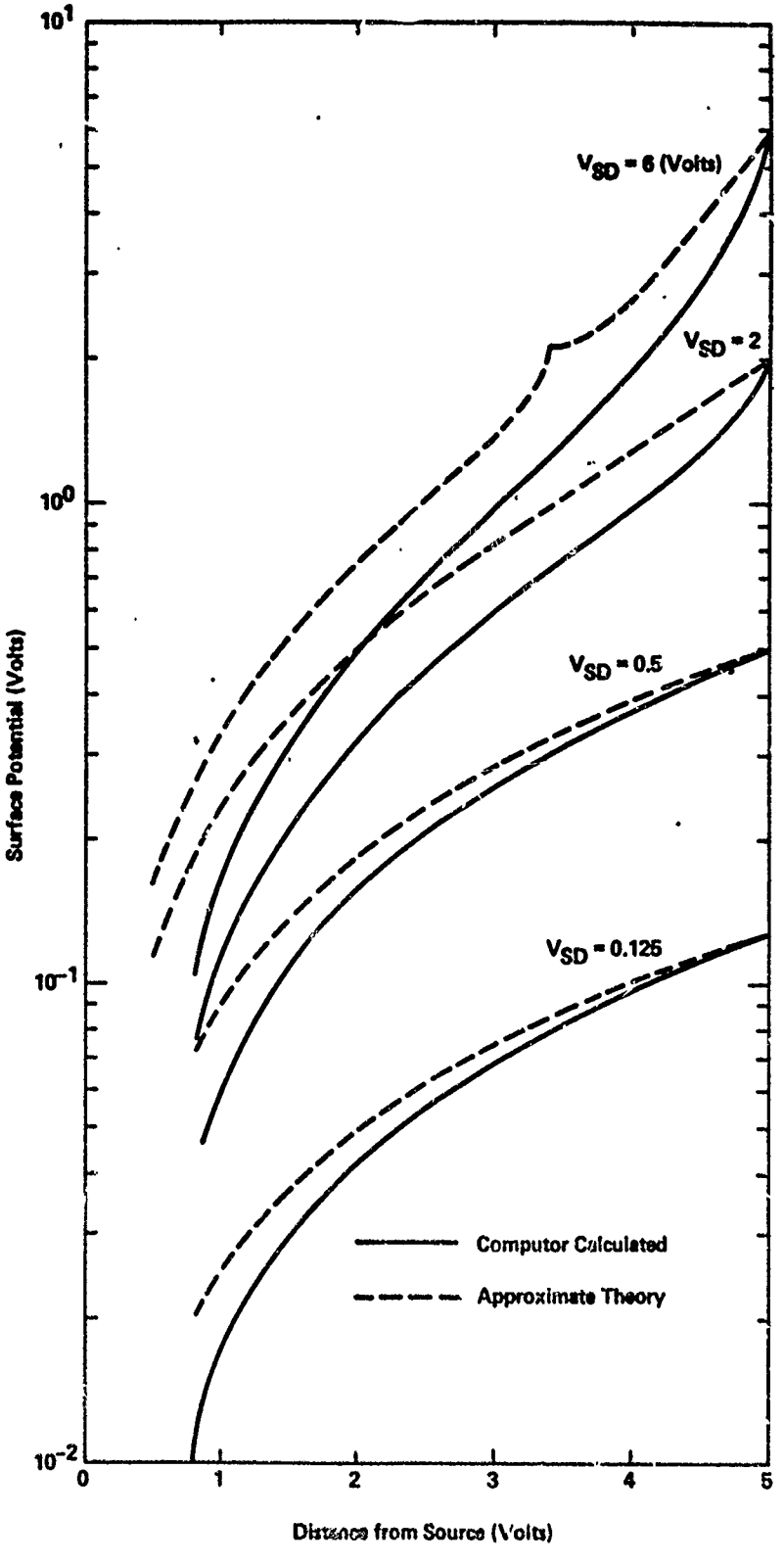

Fig. 44. Calculated potential distribution along the oxide-silicon interface in a 5.0 μm MOSFET (V_{SG} = 3.0 volts).

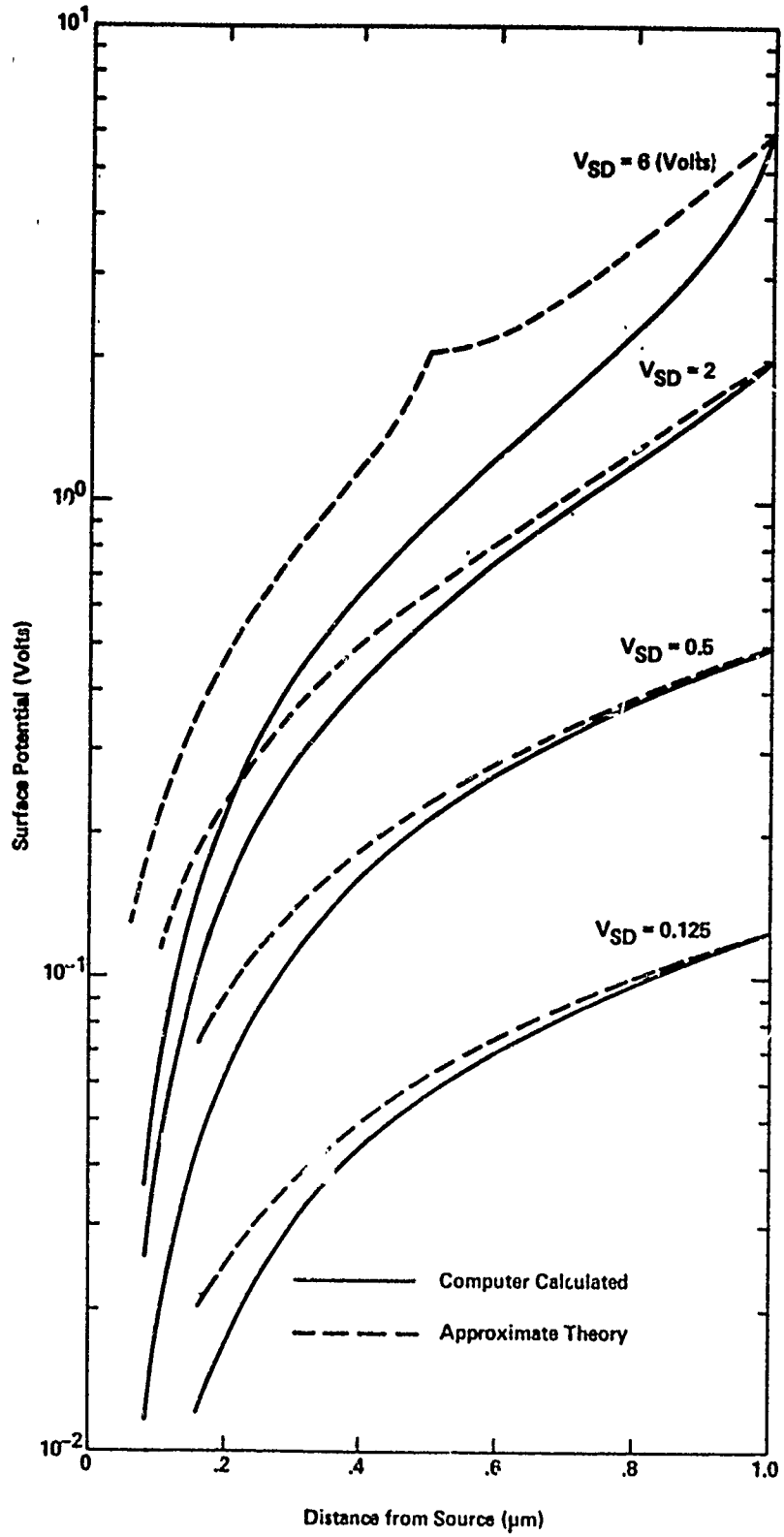

Fig. 45. Calculated potential distribution along the oxide-silicon interface in a 1.0 μm MOSFET (V_{SG} = 3.0 volts).

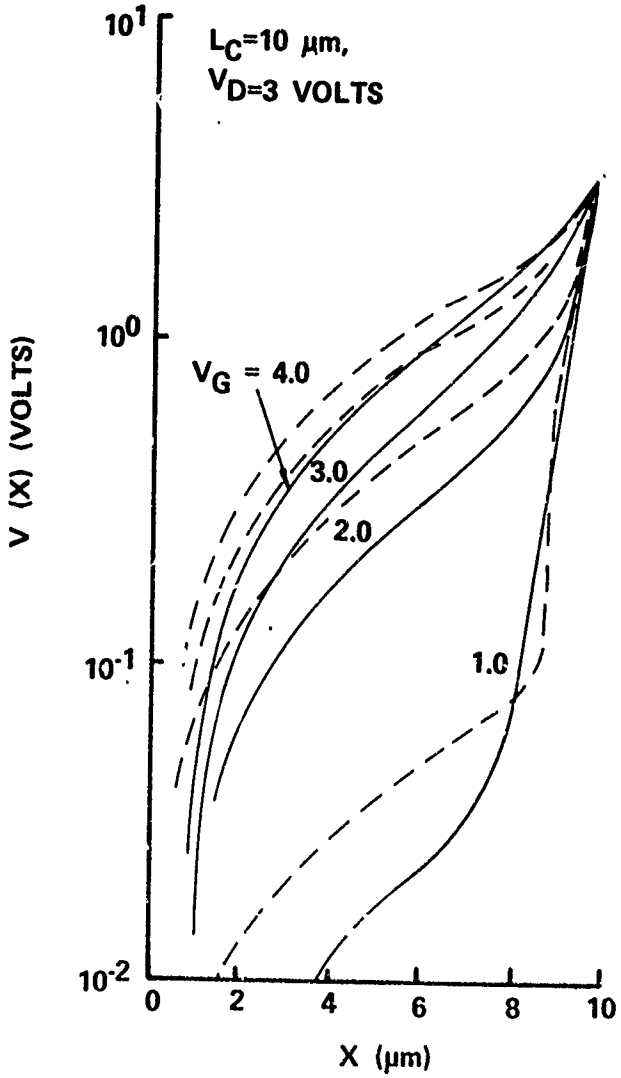

Fig. 46. Calculated potential distribution along the oxide-silicon interface in a 10.0 μm MOSFET (V_{SD} = 3.0 volts).

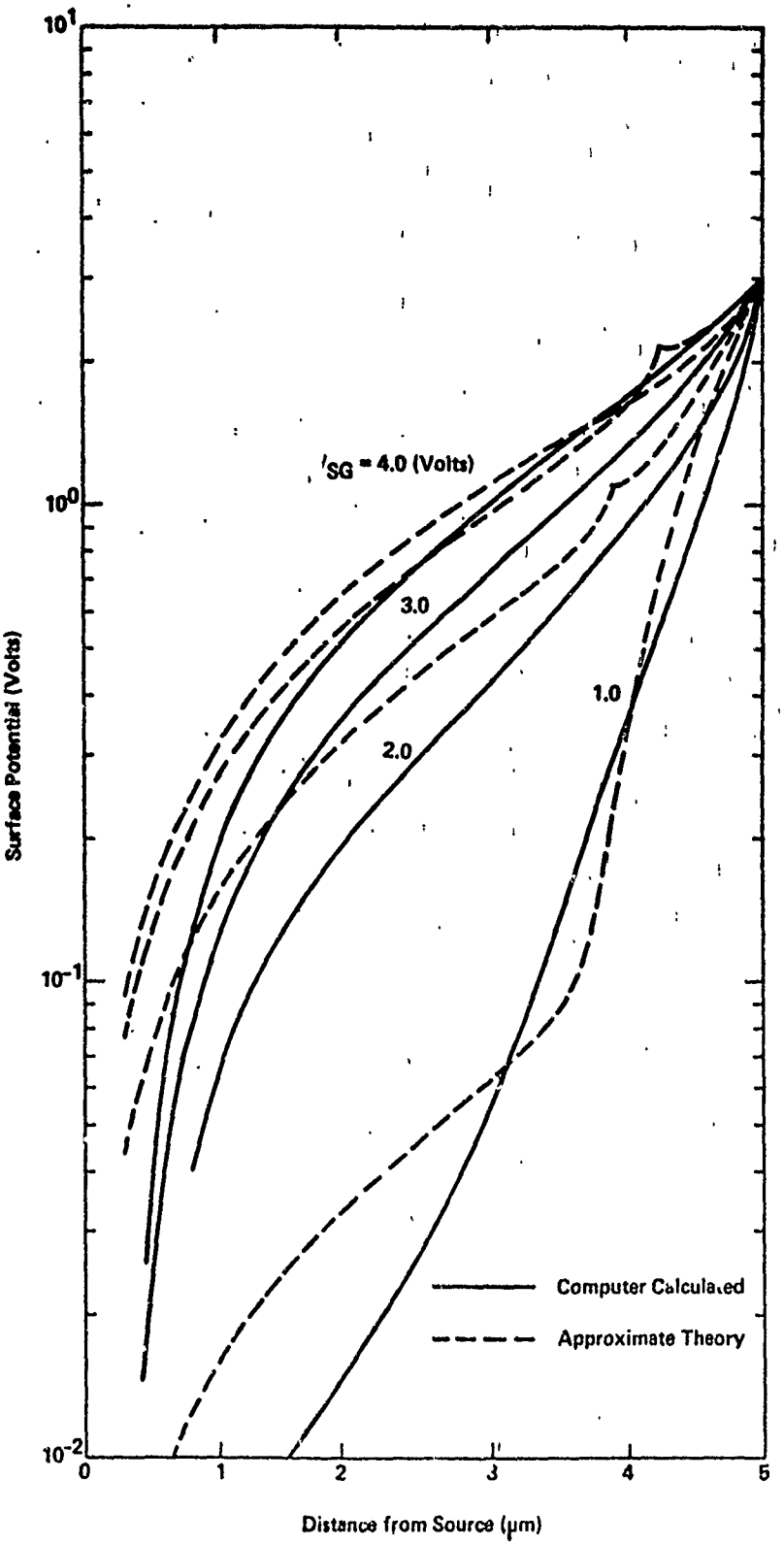

Fig. 47. Calculated potential distribution along the oxide-silicon interface in a 5.0 μm MOSFET (V_{SD} = 3.0 volts).

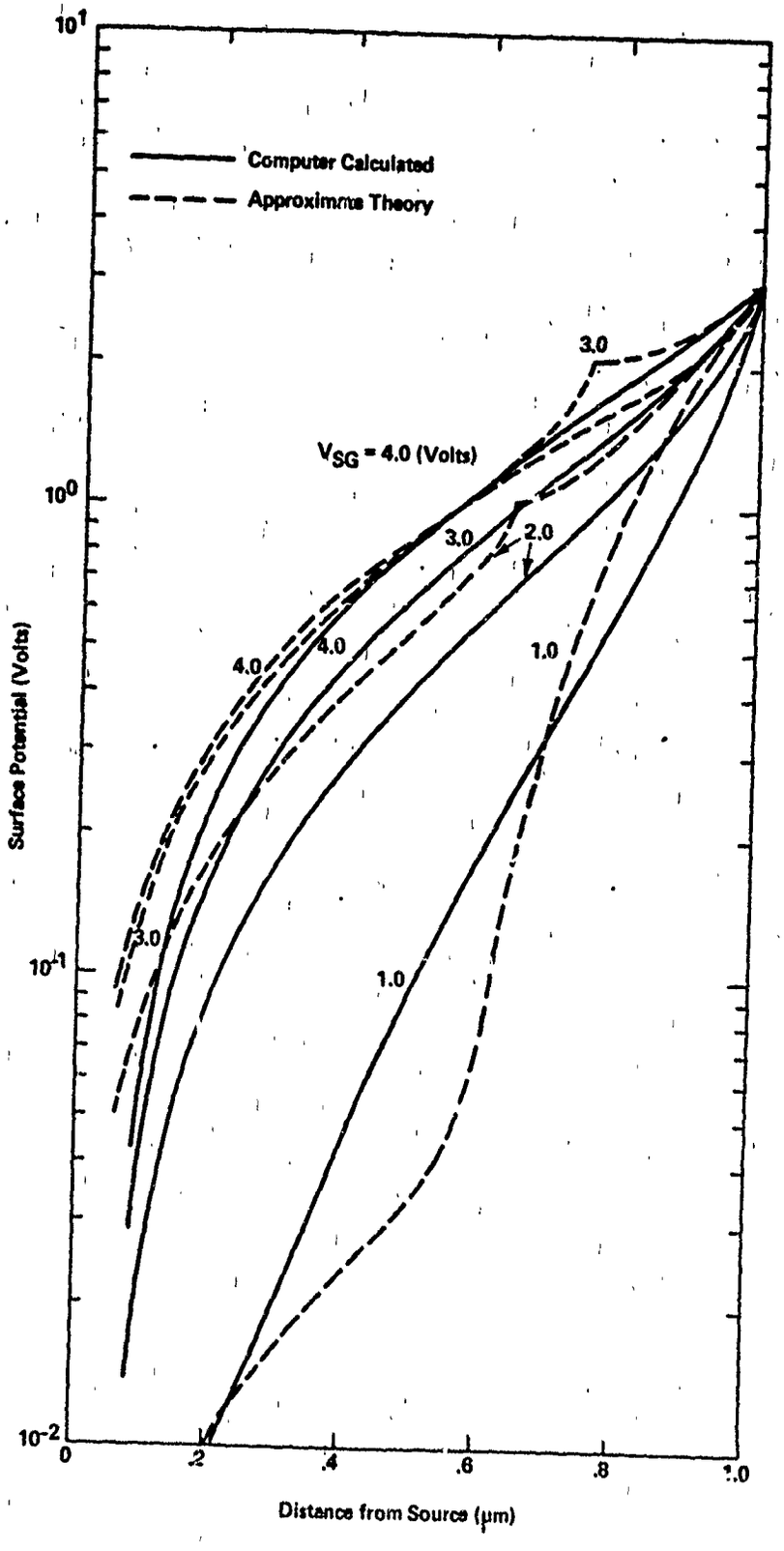

Fig. 48. Calculated potential distribution along the oxide-silicon interface in a 1.0 μm MOSFET (V_{SD} = 3.0 volts).

distribution. At best, this one-dimensional theory provides only a first-order type of approximation, and should be used with caution in the evaluation of IGFET structures.

Similarly, Figs. 46-48 present a comparison between these one-dimensional and two-dimensional calculations when the source-drain voltage is held constant (V_{SD} = 3.0 volts) and the source-gate voltage is changed from 1.0 to 4.0 volts. At small values of gate voltage, the appreciable error in this approximate theory appears to be attributable to a poor characterization of the channel potential distribution near the drain end of the structure. As previously shown, an increase of gate voltage produces a significant increase of drain junction penetration along the source-drain channel; a situation that has not been approximated in this one-dimensional theory.

2.7 References

1. H. K. J. Ihantola, *Stanford Electronics Lab.*, Report No. 1661-1, Sept. 1961.
2. W. Shockley, *Proc. IRE*, 40, 1365 (1952)
3. J. E. Johnson, *Solid-State Electronics*, 7, 861 (1964)
4. T. L. Chiu and C. T. Sah, *Solid-State Electronics*, 11, 1149 (1968)
5. H. C. Pao and C. T. Sah, *Solid-State Electronics*, 9, 927 (1966)
6. M. B. Barron, *Stanford Electronics Lab. Report* SEL-69-069
7. D. Vandorpe and N. H. Xuong, *Electronics Letters*, 7, 47 (1971)
8. D. P. Kennedy, *Mathematical Simulation of the Effects of Ionizing Radiation on Semiconductors*, Scientific Report #1 AFCRL Contract No. F 19628-70-C-0098 April 1971
9. W. Van Roosbroeck, *Bell Syst. Tech. J.*, 29, 560 (1950)
10. C. T. Sah (private communication)
11. G. F. Neumark, *Solid-State Electronics* 10, 169 (1967)
12. A. Bar-Lev and S. Margalit, *Solid-State Electronics*, 13, 1541 (1970
13. F. F. Fang and A. B. Fowler, *Jour. Appl. Phys.*, 41, 1825 (1970)
14. E. J. Ryder, *Phys. Rev.*, 90, 766 (1953)
15. A. B. Grebene and S. K. Ghandhi, *1968 Int. Solid-State Circuit Conf. Digest*, Philadelphia (1968)
16. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. & Dev.*, 14, 95 (1970)
17. L. Gregor, IBM Labs. (private communication)
18. Robert H. Crawford, *"MOSFET in Circuit Design"*, McGraw Hill, N.Y. (1967)
19. F. Fang, IBM Labs. (private communication)

CHAPTER III

Bipolar Transistor Investigations

3.0 Introduction

Important questions remain concerning the physical mechanisms involved in bipolar transistor operation. In particular, there is extensive technical literature on the theory of current gain in bipolar transistor operation (both high frequency and low frequency), although a detailed comparison with experiment shows this theory has many deficiencies. Most comparisons between theory and experiment show the measured f_t of a transistor is two-thirds to one-half its calculated value. This difference between experiment and theory appears to arise from an inherent inadequacy of our mathematical models: the theory of bipolar transistors is based upon one-dimensional and pseudo two-dimensional models that do not take into consideration detailed two-dimensional mechanisms encountered in this semiconductor device.

Numerous workers have conducted a pseudo two-dimensional analysis of bipolar transistor operation by coupling together several one-dimensional solutions through either a base region resistor[1] or through more complicated base region mechanisms.[2] This technique is an obvious attempt to overcome computational difficulties inherent in any truly two-dimensional numerical solution of this problem. Between six and twelve one-dimensional arrays are used to describe the intrinsic region of a bipolar transistor, and these arrays are coupled together at three locations: in the emitter region, in the collector region, and through a base region element. This technique has the advantage of computational simplicity, yet there is little information available to establish the errors introduced by such an approximation method.

It is for these reasons we have undertaken the task of developing a rigorous two-dimensional solution of the bipolar transistor problem. The complications associated with running this program are recognized to render it inapplicable for the solution of engineering problems on a routine basis. The intent here is to use this program in an investigation of the mechanisms involved in transistor operation and, thereby,

obtain a verification of the more approximate computational techniques presently in use.

Our computer program utilizes a matrix of 2800 points that are spatially distributed to reduce errors associated with approximating a transistor in this fashion. An important difficulty is the computation time required to attain a single solution; this computation time varies significantly with the assumed operating level of the transistor. At low levels of emitter current only a small amount of base region conductivity modulation is realized and, as a consequence, a solution is attained in a few minutes (360/Mod 91). In contrast, at large values of emitter current the structure exhibits substantial conductivity modulation; this situation can increase the computation time to in excess of 4.0 hours.

Although the computation time is excessive, using this bipolar transistor program, a long computation time is not inconsistent with previous experience. All programs developed during the past few years undergo several phases of development: the initial prototype always requires an excessive amount of computation time and, after further program development, the required computation time decreases significantly. For this reason, before we can use this computer program for an extensive investigation of bipolar transistor operation, further work is necessary to reduce the computation time to an economically practical level.

The present discussion is intended to outline qualitative observations based upon initial test calculations using this computer program. One sequence of calculations was conducted for a typical bipolar transistor structure, throughout a range of emitter current densities up to about $3.8 \times 10^4 \text{amps/cm}^2$. The principal purpose for this effort was to establish the operating characteristics of the program and, as a secondary issue, to qualitatively study some concepts of transistor operation that are in question. In particular, an attempt was made to qualitatively establish the degree of emitter current crowding obtained in a bipolar transistor at large emitter current densities. In conjunction with these calculations, a qualitative study was also made to establish the mechanisms of current gain fall-off at large values of emitter current density.

3.1 Analysis

It has been shown that the hole and electron distributions in semiconductor material are described by the equations[9]

$$\text{a)} \quad \text{div grad } \psi = \frac{-q}{\kappa\varepsilon_0}(N - n + p)$$

$$\text{b)} \quad \vec{J}_p = -qD_p \text{ grad } p - q\mu_p p \text{ grad } \psi$$

$$\text{c)} \quad \vec{J}_n = qD_n \text{ grad } n - q\mu_n n \text{ grad } \psi \qquad (1)$$

$$\text{d)} \quad \text{div } \vec{J}_p = qR_p$$

$$\text{e)} \quad \text{div } \vec{J}_n = qR_n$$

$$\text{f)} \quad \vec{J}_T = \vec{J}_p + \vec{J}_n$$

when it is assumed that no trapping mechanisms exist within the structures under consideration. Finite difference methods are used to numerically solve this system of equations; details of these computational methods are outlined in Chapter IV.

As in all mathematical investigations of bipolar transistor operation, little information is available concerning the minority carrier lifetimes within the base and emitter region. Therefore, for this initial sequence of calculations we have assumed the hole (emitter) and electron (base) minority lifetimes are identical, and the magnitude of these lifetimes was adjusted to yield a calculated current gain of about 200; this was accomplished with an assumed lifetime of 10 nsec.

Carrier mobilities in these calculations are dependent upon local impurity atom densities, local electric field, and direction.[4] The characterization of this physical parameter at large values of electric field is consistent with published information on this topic.[5] Specific information on the methods used to describe this minority carrier mobility is outlined in Chapter IV.

Throughout this mathematical investigation, particular emphasis is placed upon the selection of boundary conditions that do not introduce errors in the calculated results. In particular, the boundary

conditions used in this analysis approximate physical and electrical characteristics at the outer periphery of this semiconductor structure, rather than along internal boundaries established by its physical or electrical properties. The exposed semiconductor surface is assumed to be an ideal electrical insulator; i.e., no electric current is permitted normal to these bounding surfaces. The ohmic contacts are approximated by equipotential surfaces that are charge neutral; where appropriate, these ohmic contacts have been located sufficiently far from the active regions of this device to have no influence upon its electrical properties.

Figure 49 shows the analytical model used in these two-dimensional transistor calculations. In this model we have eliminated a large portion of the extrinsic base region and, thereby, concentrated the available matrix points in the intrinsic portion of the structure. It is emphasized, the base ohmic contact location in Fig. 49 is a sufficient distance from the emitter junction so that this contact has no influence upon the fundamental mechanisms of transistor operation.

The base region in this structure is assumed to result from a two-step diffusion[6] into n-type semiconductor material of homogeneous background doping (10^{16}atoms/cm^3). For this model the boron profile is a one-dimensional Gaussian distribution with a surface concentration of 10^{19}atoms/cm^3. The emitter distribution has been calculated using a 5 μm diffusion mask opening; the profile calculation is based upon a previous publication on this topic.[7] It is assumed this emitter junction arises from a phosphorus surface concentration of 5×10^{20}atoms/cm that is diffused for a sufficient period to produce an emitter junction depth of 1.5 μm. All calculations are based upon an assumed emitter junction length of 10.0 μm.

3.2 Current Gain in a Bipolar Transistor

At relatively small values of emitter current, it is experimentally observed that the current gain of a transistor increases with an increase of emitter current. In contrast, at large values of current this gain is observed to decrease with an increase of emitter current. These experimental observations are consistent with our calculations of current gain

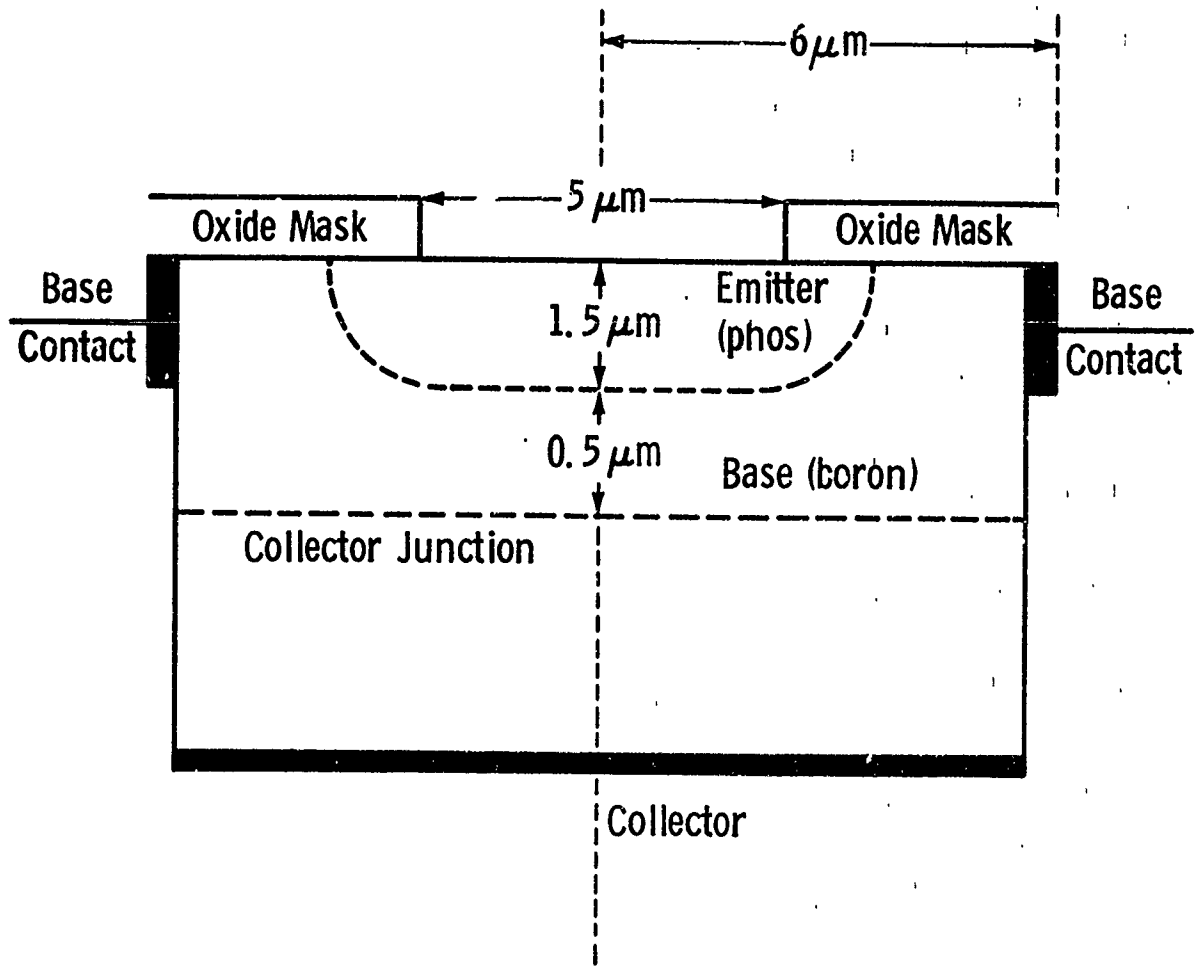

Fig. 49. Analytical model for a bipolar transistor.

(Fig. 50), although it will be shown that traditional explanations for this situation are inconsistent with conclusions drawn from our two-dimensional analysis.

In the low current region of Fig. 50, an increase of current gain (with an increase of emitter current) is traditionally attributed to ambipolar diffusion of minority carriers across a transistor base region. This mechanism was first proposed by W. M. Webster[8] (in 1954) and has been generally accepted throughout the semiconductor industry. Briefly, it is presumed that conductivity modulation within the transistor base region produces an electric field that enhances minority carrier transport from emitter to collector. At large values of conductivity modulation, it can be shown this electric field enhancement is equivalent to doubling the base region minority carrier diffusion coefficient. An increase of base region transport efficiency (due to ambipolar diffusion mechanisms) is presumed to produce the observed increase of transistor gain, with an increase of emitter current.

Although ambipolar diffusion mechanisms do take place in the transistor base region, rigorous calculations show that this mechanism does not necessarily produce the gain vs. emitter current characteristics observed in diffused transistor operation. Figure 51 illustrates the one-dimensional impurity atom distribution in a typical diffused bipolar transistor. Included in this illustration are arrows (marked J) indicating the component of minority carrier flux (electrons) attributable to drift mechanisms within the transistor base region, at small levels of base region conductivity modulation. These arrows show the base region electric field in a diffused transistor does not necessarily enhance minority carrier transport between the emitter and collector, as presumed in the theory of drift transistors.[9] Near the emitter junction, compensation of base region impurity atoms by emitter region impurity atoms yields a "built-in" electric field that opposes minority carrier transport from emitter to collector; this opposition produces a reduction of base region transport efficiency.

A relatively small impurity atom density exists throughout this region of reverse "built-in" electric field; therefore, this region is highly susceptible to conductivity modulation. Calculations indicate

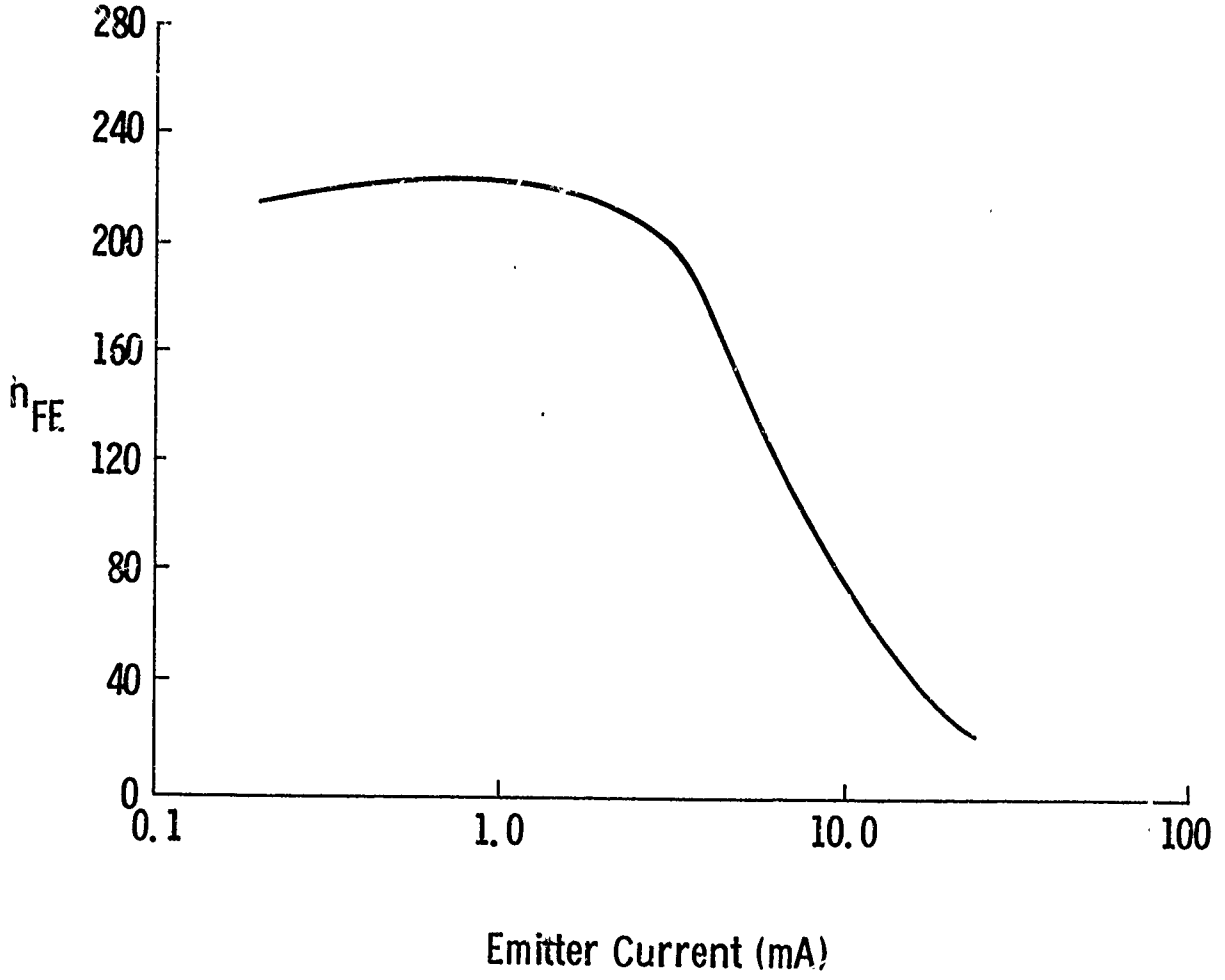

Fig. 50. Calculated current gain for the bipolar transistor shown in Fig. 49.

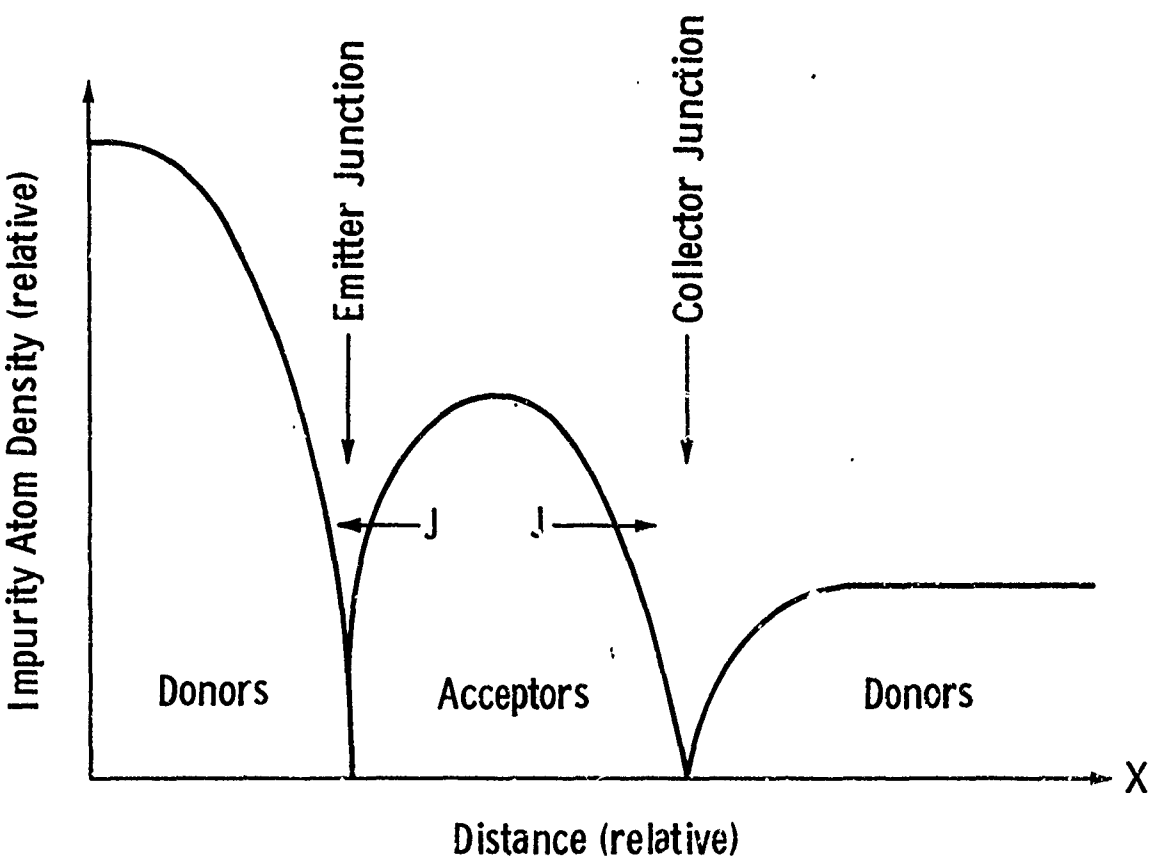

Fig. 51. Illustrative one-dimensional impurity atom profile for a diffused bipolar transistor.

that at small values of emitter current conductivity modulation essentially eliminates this reverse "built-in" field and, therefore, increases the base region transport efficiency. Thereby, an increase of transistor gain, with an increase of emitter current, is obtained at relatively small values of emitter current.

Traditional theory predicts that at large emitter current densities bipolar transistors will exhibit a decrease of current gain, with an increase of emitter current, although this theory is based upon a one-dimensional analysis of the problem. Our calculations of current gain (Fig. 50) are consistent with this prediction, yet the mechanisms producing a decrease of gain are essentially two-dimensional. From these calculations, it is suggested that a substantial modification is necessary in the traditional theory of bipolar transistor operation before this theory is entirely consistent with the physical mechanisms involved in device operation.

Figure 52 shows the calculated base region transport efficiency (β_B) and emitter region injection efficiency (β_γ) for this analytical model. These parameters (β_B, β_γ) represent the current gain that would be exhibited by this transistor if the individual parameter under consideration represented the only source of gain limitation. For example, β_B represents the current gain if the transistor contained an ideal emitter junction (γ=1). Similarly, β_γ represents the transistor current gain assuming no minority carrier recombination mechanisms within the base region.

From Fig. 52, both emitter injection efficiency (β_γ) and base region transport efficiency (β_B) undergo a substantial decrease at large values of emitter current. Traditional theory would attribute this loss of base transport efficiency (β_B) to the so-called "Kirk-effect": collector junction displacement (and, hence, base region widening) due to a large minority carrier density within the collector junction space-charge layer.[10] In addition, this traditional theory would attribute the loss of injection efficiency (β_γ) to base region conductivity modulation. Although any two-dimensional structure must, indeed, exhibit the mechanisms suggested by this traditional

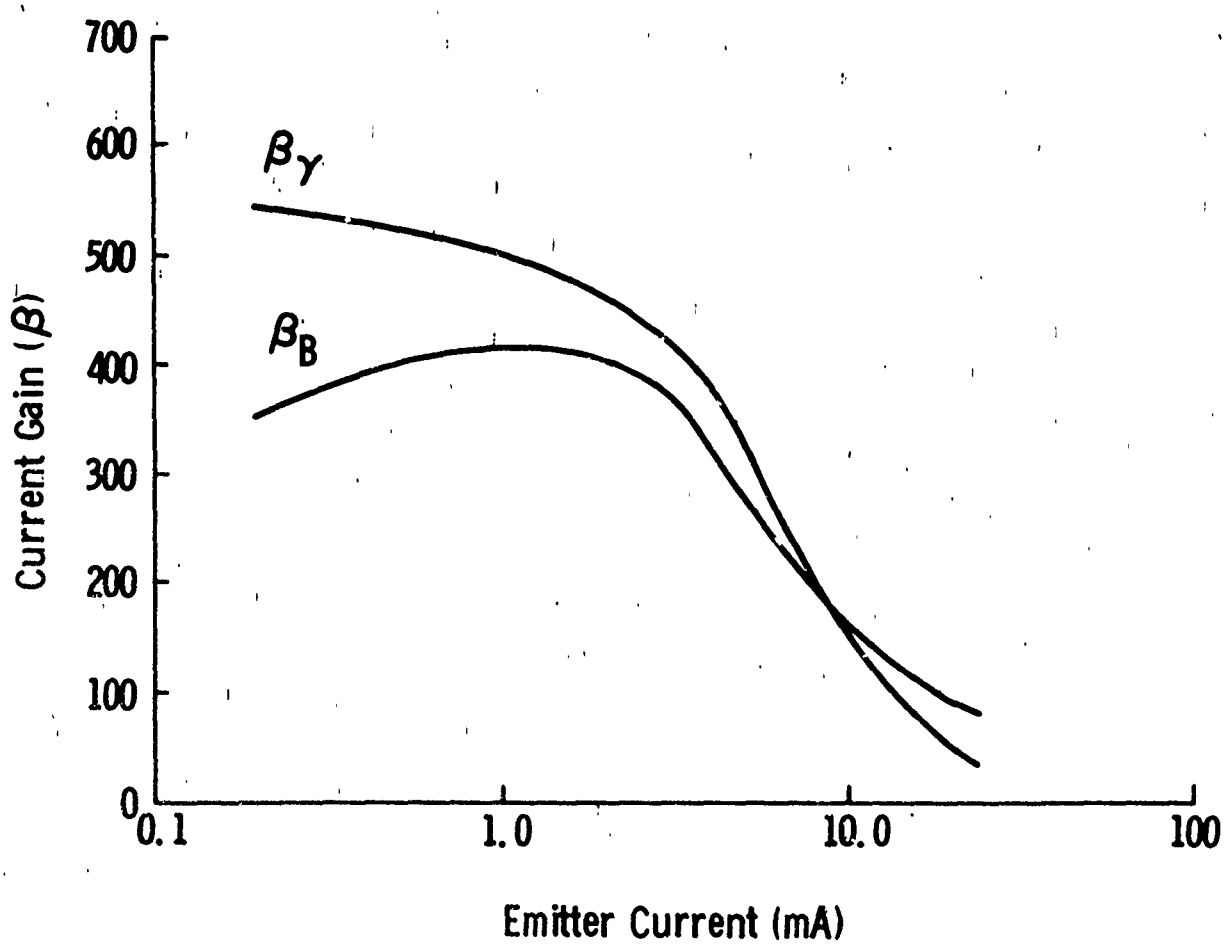

Fig. 52. Calculated emitter transport and base region transport efficiencies for a bipolar transistor (Fig. 49).

theory, there is no reason to assume that these mechanisms produce the observed current gain characteristics of a diffused bipolar transistor.

Calculations indicate that base region widening, due to minority carriers within the collector space-charge layer (Kirk-effect), is of little consequence in this particular semiconductor device. Instead, the calculated decrease of base region transport efficiency (β_B) appears to result from side-wall injection of minority carriers by the emitter junction. This situation is illustrated in Fig. 53. In Fig. 53, a dashed line is used to designate that region of the emitter junction where the current density (due to electron injection) is greater than one-half its maximum value. At small values of total emitter current (0.19 ma) only a small amount of electron injection takes place along the emitter side-wall. In contrast, at large values of total emitter current (24 ma) substantial side-wall injection is encountered in this semiconductor device.

A consequence of this side-wall injection is an increase of effective base width, with an increase of emitter current (see Fig. 54). Carriers entering the base region from the emitter junction side-wall must traverse a longer path than carriers entering from other regions of the emitter; these side-wall injected carriers exhibit substantial base region recombination and, thereby, produce the calculated fall-off of base transport efficiency (β_B). It is emphasized, at high emitter currents the region of base region recombination differs significantly from that suggested by traditional theory of bipolar transistors. Traditional theory suggests that most base region recombination takes place under the emitter, rather than along its periphery, and this concept has been adopted in most pseudo two-dimensional investigations of device operation.[1,2]

As previously stated, it is generally assumed that the decrease of emitter injection efficiency with emitter current (β_γ in Fig. 52) is a result of base region conductivity modulation. All one-dimensional (and pseudo two-dimensional) calculations of device performance are based upon an assumption that this is the dominant mechanism producing β_γ fall-off. Although base region conductivity modulation may, indeed,

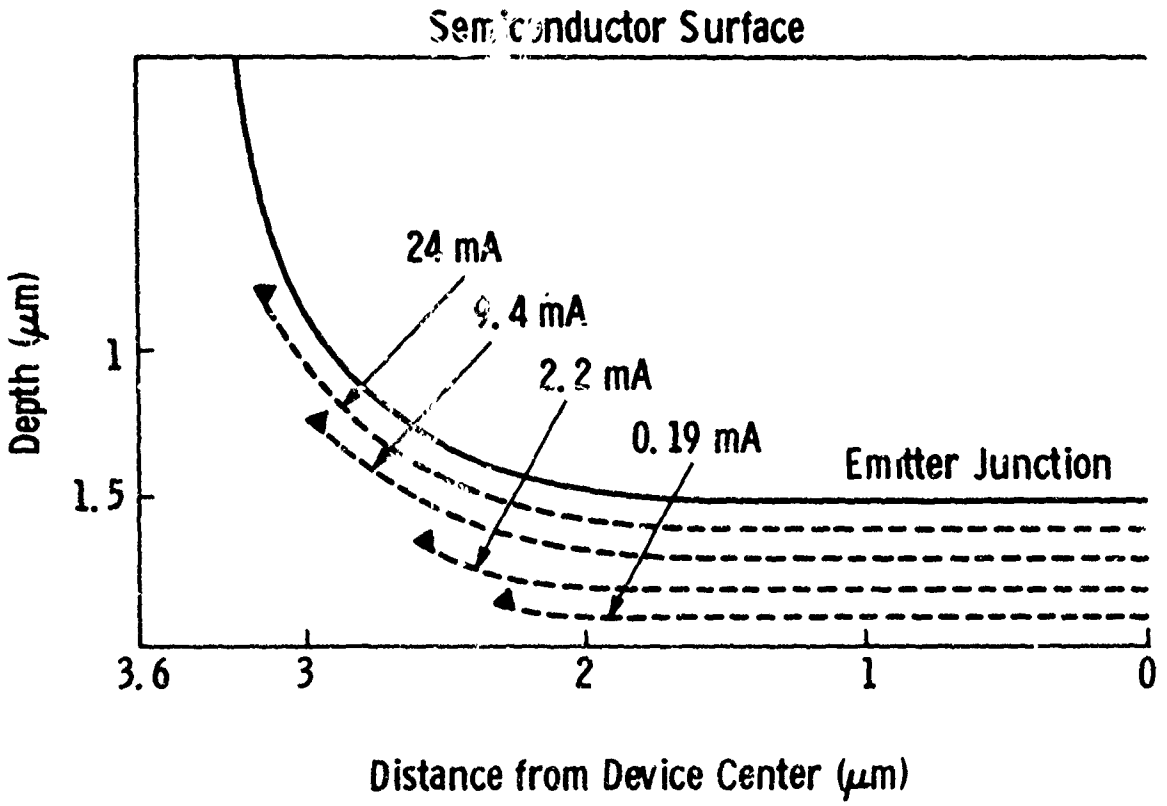

Fig. 53. Illustration of the calculated active emitter junction area with increasing emitter current

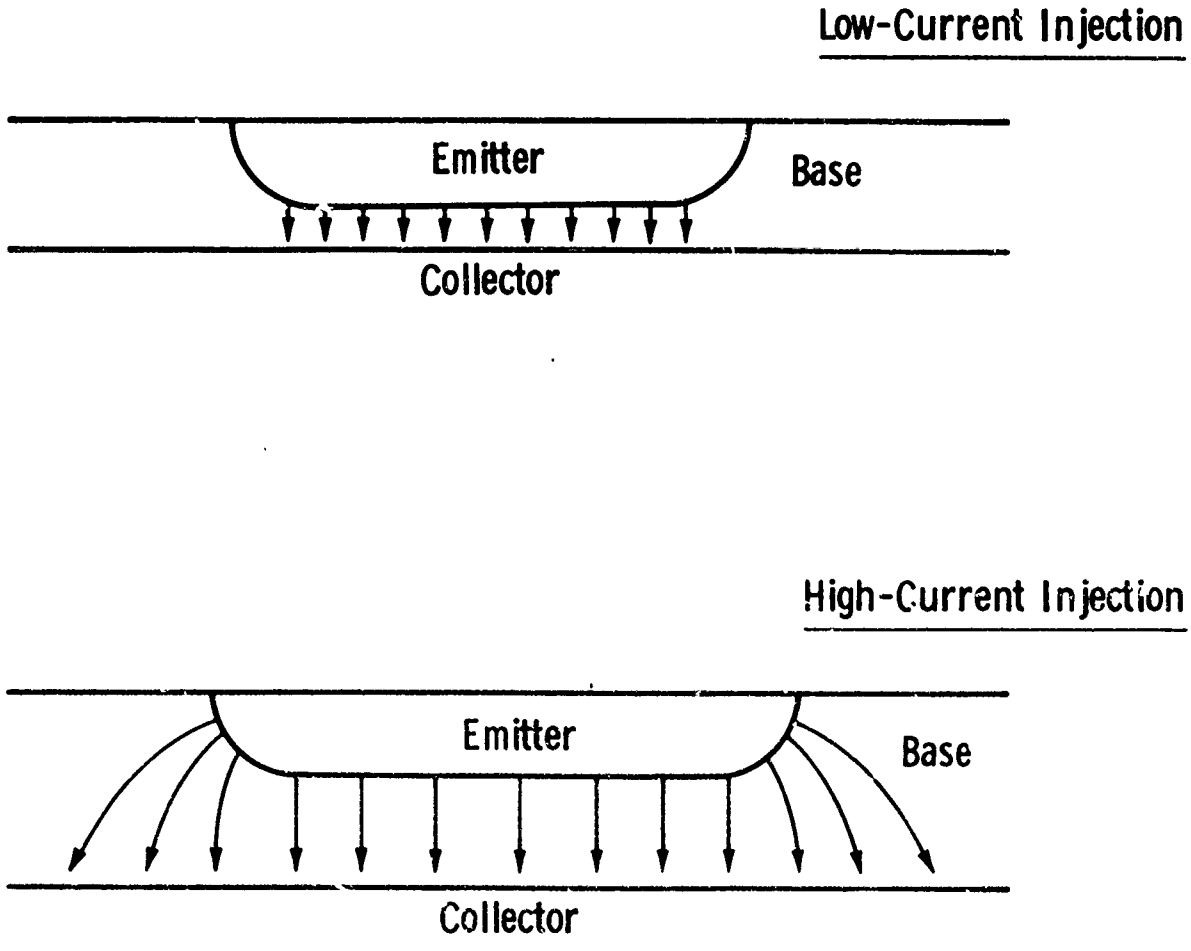

Fig. 54. Illustration of the calculated minority carrier flux distribution in a bipolar transistor.

produce a loss of emitter injection efficiency, Fig. 55 shows another mechanism that has a more significant influence upon the electrical characteristics of this particular bipolar transistor.

Figure 55 shows a sketch of this two-dimensional emitter junction upon which is drawn dashed lines to indicate the region of maximum back injection (holes) into the emitter. At each value of total emitter current, a given dashed line bounds that portion of the junction exhibiting a back injection (holes injected into the emitter region) greater than one-half its maximum value. At a small value of emitter current, little back injection takes place at the emitter side-wall. In contrast, at large values of emitter current (14 ma) almost all the loss of injection efficiency is attributable to side-wall injection of holes into the emitter region.

A consequence of this situation is that the loss of emitter injection efficiency is attributable to side-wall injection into the emitter region, rather than from the intrinsic region of this structure (Fig. 56). It is emphasized, at large values of emitter current the distribution of base region current, due to back injection, differs significantly from the distribution suggested by elementary theory. At large values of emitter current the base region flux, due to this mechanism, is concentrated near the emitter periphery, rather than between the emitter and collector junctions.

The foregoing represent some initial conclusions drawn from this two-dimensional analysis. These calculations show the side-wall of an emitter junction has an important influence upon the electrical properties of this semiconductor device. Further study of the computed results is required before we can quantitatively establish the mechanisms producing side-wall injection and, thereafter, generate an elementary (or engineering) theory that is consistent with the physics of bipolar transistor operation.

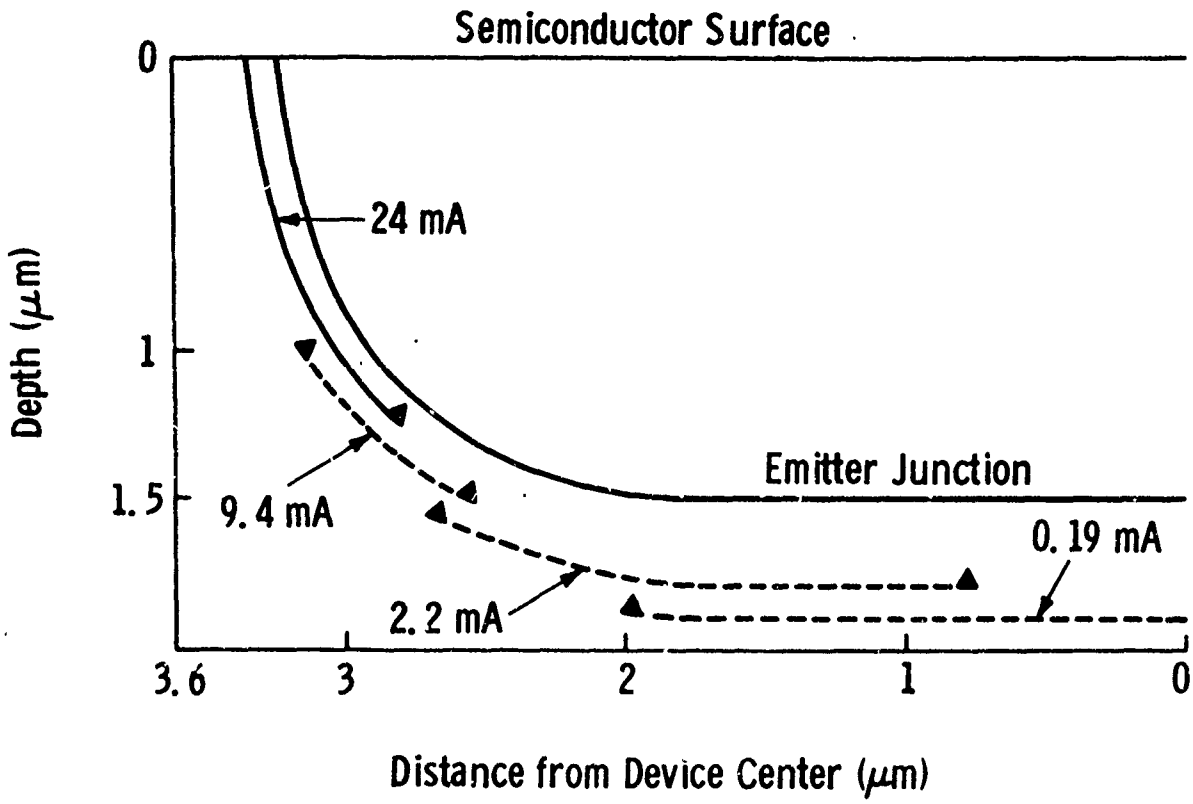

Fig. 55. Illustration of the calculated region of back injection by the emitter junction with increasing emitter current

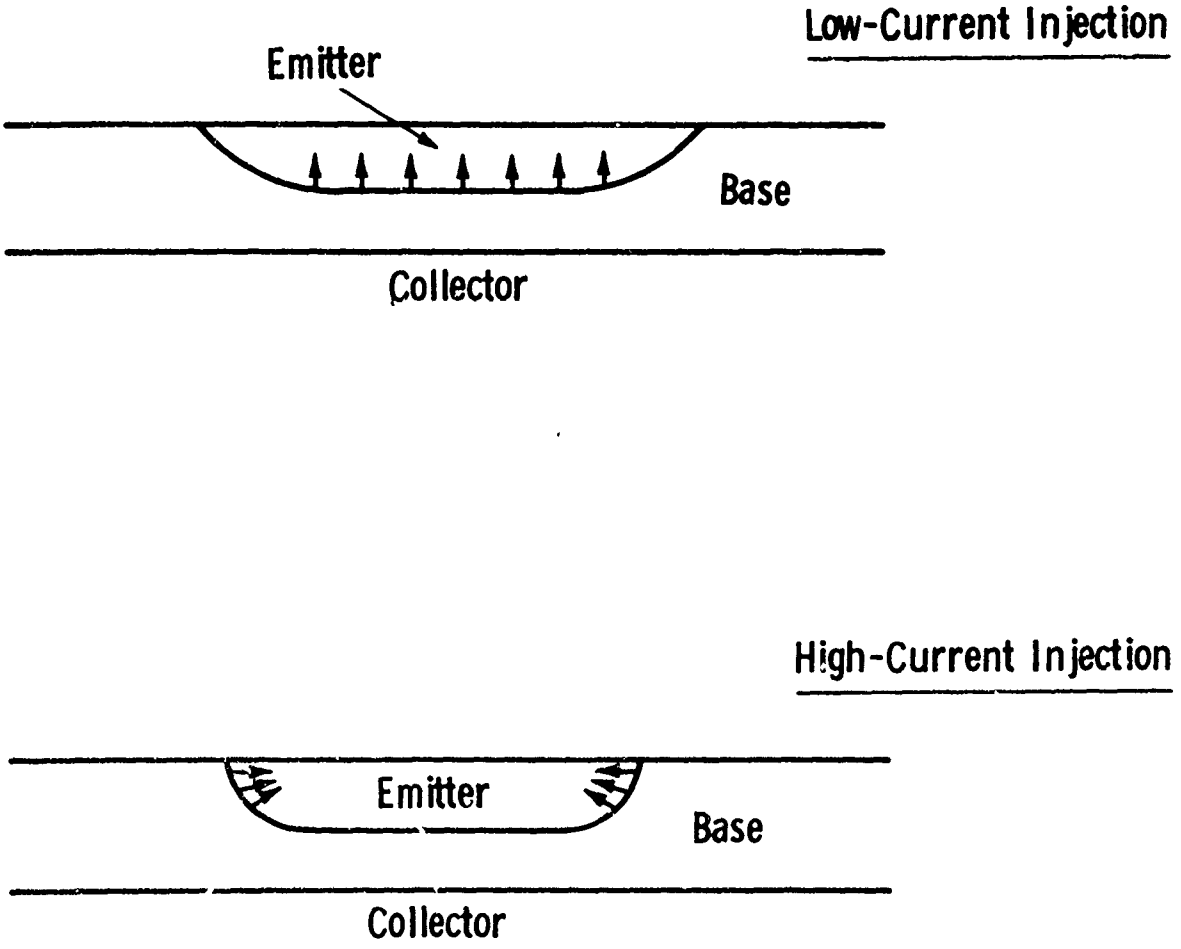

Fig. 56. Illustration of the calculated back injection of a diffused emitter in a bipolar transistor.

3.3 Emitter Current Crowding

Traditional theory of transistor operation is based upon an assumption that all base region majority carrier current is attributable to mechanisms encountered between the emitter and collector junctions (base region recombination and back injection by emitter junction), as illustrated in Fig. 57. An important consequence of this assumption is a longitudinal variation of emitter biasing voltage that produces a substantial increase of emitter current density near the emitter edge (current crowding). It is shown here that emitter current crowding is of minor consequence in this particular mathematical model; present pseudo two-dimensional analysis of this same structure indicates substantial current crowding.

Figure 58 shows the calculated emitter crowding in this semiconductor device, as a function of total emitter current. From this sequence of calculations, at a total emitter current of 13 ma, this structure exhibits a current crowding of only 0.8: ratio of current density at center of emitter to current density at emitter edge. In contrast, a pseudo two-dimensional calculation would show a substantially greater current crowding ratio at this same total emitter current. The reasons for this disagreement are evident in view of the current gain mechanisms outlined in section 3.3.

Figure 59 illustrates the path of majority carrier flux in this transistor base region. At small values of emitter current this majority carrier flux appears under the emitter junction, although its magnitude is insufficient to produce a significant amount of emitter crowding. At large values of emitter current, base region majority carrier flux is directed predominantly toward the emitter side-wall and recombination centers located outside the intrinsic base region; this base region flux cannot influence the longitudinal distribution of emitter bias and, hence, cannot produce current crowding. As a consequence, a rigorous two-dimensional solution of the bipolar transistor problem shows substantially less emitter current crowding than is suggested by the present technical literature.

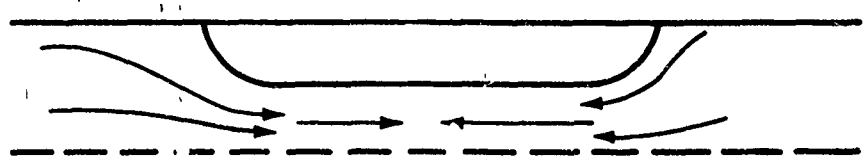

Fig. 57. Illustration of the assumed majority carrier flux in a transistor base region (traditional theory).

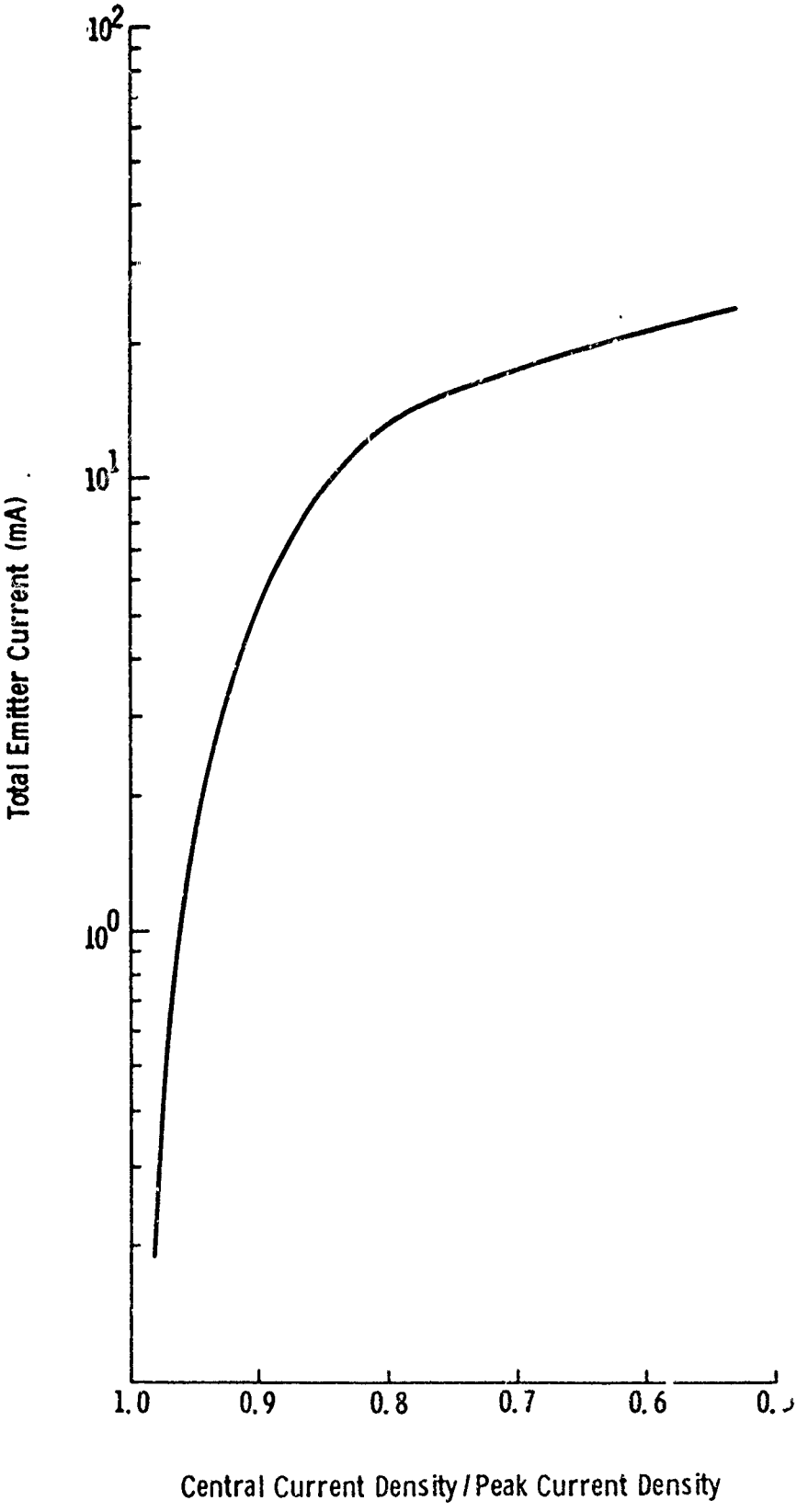

Fig. 58. Calculated ratio of current density at the center of emitter and at emitter edge.

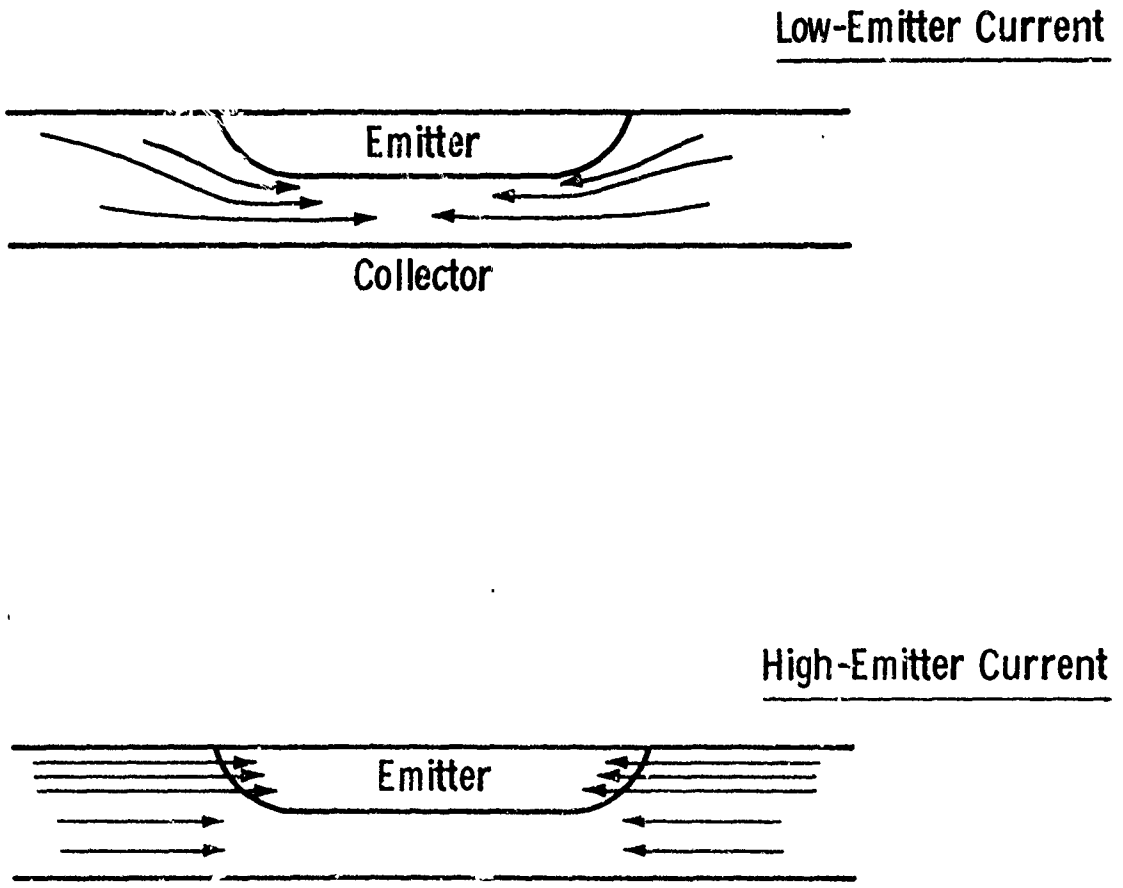

Fig. 59. Illustration of base region majority carrier flux in a bipolar transistor.

3.4 References

1. H. N. Ghosh, *IEEE Trans.*, ED-12, 513 (1965)
2. B. V. Gokhale, *IEEE Trans.*, ED-17, 594 (1970)
3. W. Van Roosbroeck, *Bell Syst. Tech. J.*, 29, 560 (1950)
4. D. M. Caughey and R. F. Thomas, *Proc. IEEE (Letters)*, 55, 2192 (1967)
5. E. J. Ryder, *Phys. Rev.*, 90, 766 (1953)
6. D. P. Kennedy and P. C. Murley, *Proc. IEEE*, 51, 620 (1964)
7. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. and Dev.*, 9, 179 (1965)
8. W. M. Webster, *Proc. IRE*, 42, 914 (1954)
9. H. Krömer, *Naturwissenschaften*, 40, 578 (1953)
10. C. T. Kirk, *IRE Trans.*, ED-9, 164 (1962)

CHAPTER IV

Computational Methods

4.0 Introduction

In this chapter we describe the two-dimensional mathematical methods used throughout this investigation. Our methods are based on a well-known system of partial differential equations describing steady-state carrier distributions in semiconductors.[1] The geometry and boundary conditions adopted to approximate the structures under consideration are intended to be consistent with idealized devices. Finite difference techniques are applied to this system of differential equations; an algorithm is presented for computing solutions for these equations as applied to an IGFET and bipolar transistor.

Since several papers have appeared on the use of difference schemes for semiconductor problems, this discussion will be limited to the individual details of the present method. However, some of our results do not agree with previously published computations on IGFET structures;[2,3] we therefore present enough details of the computations to justify our results.

Our analytical model for an IGFET is shown in Fig. 60. This structure is divided into two regions: first, the substrate (R_1) and, second, the oxide layer (R_2). The electrostatic potential function is defined in the region $R=R_1 \cup R_2$; the carrier concentrations, current densities, etc., are defined only in the substrate region R_1.

A number of approximations have been introduced into this analysis, to simplify the computations and the interpretation of the computed results. A n-channel device is considered, in which hole current and recombination are neglected.[4] We also assume abrupt junctions at the source and drain, a uniform nondegenerate substrate doping, and an ideal charge-free substrate-oxide interface. With these simplifications, the mathematical problem considered here is reduced to the simultaneous solution of the Poisson equation and an equation of electron current continuity, each with specified boundary conditions.

In section 4.1, the statement of this boundary value problem is made specific, and notation is presented. The employed computation scheme is described in section 4.2, and a semi-empirical algorithm for automatically calculating a suitable sequence of relaxation factors is

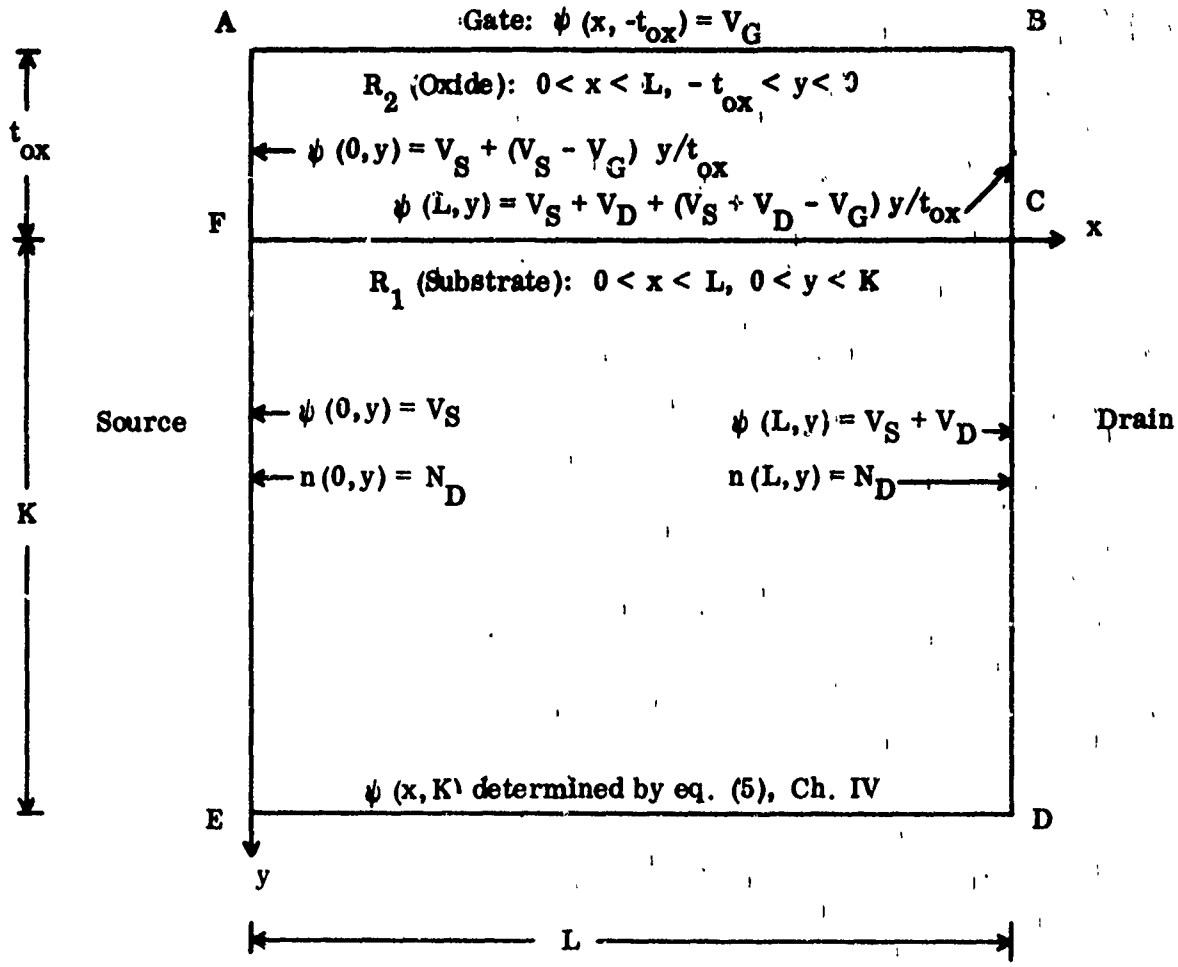

Fig. 60. Analytical model and boundary assumed conditions for an IGFET.

given in section 4.3. The difference approximations as employed are described in sections 4.4, 4.5; some additional details of the computations, and a summary of numerical experiments designed to evaluate the accuracy of the computations, are given in section 4.6.

Figure 63 illustrates the two-dimensional analytical model used throughout our investigation of bipolar transistor operation. The boundary conditions, etc., used to approximate this structure are discussed in section 4.7 and the associated computational method is described in section 4.8.

4.1 Definitions and Notation

We employ a system of units in which the intrinsic carrier density, the Boltzmann voltage, and the electronic charge have magnitude unity; letting $\psi(x,y)$, $n(x,y)$, $p(x,y)$ denote, respectively, the electrostatic potential, electron concentration, and hole concentration functions, the Poisson equation becomes

$$\text{a)} \quad \varepsilon_s \Delta\psi - N_A + p - n = 0, \quad (x,y)\varepsilon R_1; \tag{1}$$

$$\text{b)} \quad \Delta\psi = 0, \quad (x,y)\varepsilon R_2;$$

where ε_s, N_A are the substrate dielectric constant and doping, respectively. Across the substrate-oxide interface, we require ψ and $\varepsilon\partial\psi/\partial y$ continuous; the other boundary conditions are shown in Fig. 60. Similarly, the equation of electron current continuity may be written

$$\nabla(\mu e^{\psi}\nabla(e^{-\psi}n)) = 0, \quad (x,y)\varepsilon R_1 \tag{2}$$

where μ is the electron mobility, which may be position dependent and anisotropic due to dependence on the local electric field. On the substrate boundary, either n is specified (on the lines CD and EF in Fig. 60) or the normal component of the electron current density vanishes. In the neglect of hole current, the hole concentration can be written

$$p = e^{-\psi - V_{BG}} \tag{3}$$

where V_{BG} is the source-substrate or "back-gate" bias voltage.

In the following, we use V_S, V_G, V_D, respectively, for the built-in source voltage, the applied gate voltage, and the applied drain voltage. The electron quasi-Fermi potential is set to zero at the source.

We introduce the uniform and Dirichlet norms of functions defined in R,

$$\text{a)} \quad ||u|| = \sup_{(x,y)\in R} |u(x,y)| \; ; \tag{4}$$

$$\text{b)} \quad ||u||_1^2 = \int_{R_1}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] dxdy + \frac{\varepsilon_{ox}}{\varepsilon_s}\int_{R_2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] dxdy.$$

4.2 Computation Scheme

In this section we describe an algorithm for computing an approximation to a simultaneous solution of eqs. 1-3. The algorithm is based on an iteration scheme described previously,[5] and shown to converge globally in certain cases. Although the hypotheses for global convergence are not satisfied in this application, the iteration scheme has been found entirely satisfactory.

We first compute the electrostatic potential on the line ED of Fig. 60 from one-dimensional theory

$$\varepsilon_s \frac{\partial^2\psi(x,K)}{\partial x^2} - N_A + e^{-\psi(x,K)\cdot V_{BG}} - n(x,K) = 0, \quad x\varepsilon(0,L);$$

$$n(x,K) = \begin{cases} e^{\psi(x,K)}, & x \le L/2 \; , \\ e^{\psi(x,K)-V_D}, & x > L/2 \; ; \end{cases} \tag{5}$$

$$\psi(0,K) = V_S, \quad \psi(L,K) = V_S + V_D \; ;$$

these values are used as boundary conditions in subsequent calculations.

An initial approximation to the electrostatic potential distribution $\psi^{(o)}(x,y)$ in R is next computed from the Poisson equation,

$$\varepsilon_s \Delta\psi^{(0)} - N_A + e^{-\psi^{(0)}-V_{BG}} - n^{(0)}(x,y) = 0, \; (x,y)\varepsilon R \; ,$$

$$\Delta\psi^{(0)} = 0, \; (x,y)\varepsilon R_2 \; , \tag{6}$$

together with the appropriate boundary conditions. In (6), $n^{(o)}(x,y)$ is an electron distribution obtained from a previous computation with V_G and V_{BG} set at the desired values and V_D at a lower value than presently desired. (If necessary, a separate computation with V_D set to zero is made.)

The following iteration scheme is then used to obtain from the m-th iterate, $\psi^{(m)}(x,y)$, an improved approximation, denoted by $\psi^{(m+1)}(x,y)$ and the corresponding current and electron distribution $n^{(m)}(x,y)$. The field dependent electron mobilities in the x- and y-directions, denoted by μ_x and μ_y, are calculated from $\psi^{(m)}$; the equation of electron current continuity, eq. 2, is then solved for $n^{(m)}$ by describing the current density components in terms of a "stream function" $\theta(x,y)$ defined for $(x,y)\epsilon R_1$

$$\text{a)} \quad \mu_x e^{\psi^{(m)}} \frac{\partial}{\partial x}\left(e^{-\psi^{(m)}} n^{(m)}\right) = J^{(m)} \frac{\partial \theta^{(m)}}{\partial y} \quad ,$$

$$\text{b)} \quad \mu_y e^{\psi^{(m)}} \frac{\partial}{\partial y}\left(e^{-\psi^{(m)}} n^{(m)}\right) = J^{(m)} \frac{\partial \theta^{(m)}}{\partial x} \quad , \tag{7}$$

where $J^{(m)}$ is the m-th approximation to the total device current.

Equations 7 may be readily seen to imply the following equation for $\theta^{(m)}(x,y)$

$$\frac{\partial}{\partial x}\left(\frac{e^{-\psi^{(m)}}}{\mu_y}\frac{\partial \theta^{(m)}}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{e^{-\psi^{(m)}}}{\mu_x}\frac{\partial \theta^{(m)}}{\partial y}\right) = 0, \quad (x,y)\epsilon R_1 \quad . \tag{8}$$

The following boundary conditions for $\theta^{(m)}$ are equivalent to the prescribed boundary conditions for $n^{(m)}$, in view of the relations (7a, 7b):

$$\theta^{(m)}(x,0) = 1; \quad \theta^{(m)}(x,K) = 0; \quad \frac{\partial \theta^{(m)}}{\partial x}(y,0) = \frac{\partial \theta^{(m)}}{\partial x}(y,L) = 0 \quad . \tag{9}$$

Equations 8, 9 are solved for $\theta^{(m)}(x,y)$; the value of $J^{(m)}$ is then determined by integrating (7a) across the line FC(y=0) in Fig. 60, using the source and drain boundary conditions to obtain

$$J^{(m)} = -(1-e^{-V_D}) \Big/ \int_0^L e^{-\psi^{(m)}(x,0)} \frac{\partial \theta^{(m)}}{\partial y}(x,0) \frac{dx}{\mu_x} \quad . \tag{10}$$

Then the electron distribution $n^{(m)}$ is determined from (7a),

$$\text{a)}\quad n^{(m)}(x,y) = e^{\psi^{(m)}(x,y)} \left(1 + J^{(m)} \int_0^x e^{-\psi^{(m)}(z,y)} \frac{\partial \theta^{(m)}}{\partial y}(z,y)\, \frac{dz}{\mu_x} \right) \tag{11}$$

$$\text{b)}\quad = e^{\psi^{(m)}(x,y)} \left(e^{-V_D} - \int_x^L e^{-\psi^{(m)}(z,y)} \frac{\partial \theta^{(m)}}{\partial y}(z,y)\, \frac{dz}{\mu_x} \right) .$$

We remark that this method of solving the equation of current continuity is simply a generalization of the one-dimensional method of Gummel;[6] in this application, however, it has several advantages: the equation for the current distribution, eq. 8, can be solved in practice very efficiently, by a line iteration described below; at each step, approximations are obtained to the total device current and the relative current distribution; the use of appropriate difference approximations to (11) avoids numerical errors due to roundoff or machine overflow.

At each step, a function $v^{(m)}$ defined in R is computed, satisfying the boundary conditions for $\psi(x,y)$ and the relations

$$\begin{aligned} &\varepsilon_s \Delta v^{(m)} = N_A^+ - e^{-\psi^{(m)} - V_{BG}} - n^{(m)}, \quad (x,y) \in R_1 \ ; \\ &\Delta v^{(m)} = 0, \quad (x,y) \in R_2 \ ; \end{aligned} \tag{12}$$

and the improved approximation to $\psi(x,y)$ then obtained from

$$\psi^{(m+1)}(x,y) = \psi^{(m)}(x,y) + \alpha_m (v^{(m)}(x,y) - \psi^{(m)}(x,y)) + \beta_m (\psi^{(m)}(x,y) - \psi^{(m-1)}(x,y)), \quad (x,y) \in R \tag{13}$$

where α_m, β_m are constants evaluated below.

The iteration is terminated when the value of $||v^{(m)} - \psi^{(m)}||$ falls below a specified value.

4.3 Relaxation Factors

Suitable expressions for the constants α_m, β_m appearing in (13) are obtained by making the same linearization assumption originally described by Gummel,[6] and using the Chebyshev sequence of relaxation factors.[7]

Let $u^{(m)} = \psi - \psi^{(m)}$ be the error in the m-th iterate, and let S denote the set of functions which are twice continuously differentiable in R_1, harmonic in R_2, vanish on the boundary of R (the lines ABDEA of Fig. 1), and satisfy the interface relations prescribed for $\psi(x,y)$, i.e., $u(x,0+)=u(x,0-)$, $\varepsilon_s \frac{\partial u}{\partial y}(x,0+)=\varepsilon_{ox}\frac{\partial u}{\partial y}(x,0-)$, for all $u\varepsilon S$.

Clearly, $u^{(m)}\varepsilon S$. We now assume $u^{(m)}$ is small, and that the dominant mechanism affecting the propagation of the error $u^{(m)}$ is the exponential dependence of the carrier densities on the *local*[6,8] electrostatic potential, as shown in eqs. 3, 11. Then from the mean value theorem, we obtain the approximation

$$n^{(m)} - p^{(m)} \simeq n - p - (n+p)\, u^{(m)}, \tag{14}$$

where n,p denote the electron and hole concentrations in the desired solution.

It is convenient to expand $u^{(m)}$ in the set of eigenfunctions z_j, corresponding to eigenvalues λ_j satisfying

$$\lambda_j \varepsilon_s \Delta z_j + (n+p)\, z_j = 0, \quad (x,y)\varepsilon R_1 \; ; \quad z_j \varepsilon S \; . \tag{15}$$

The z_j form a complete set in S and satisfy the following orthogonality relation[9]

$$\int_{R_1} (n+p)\, z_j(x,y)\, z_k(x,y)\, dx\, dy = \delta_{j,k}. \tag{16}$$

Subtracting eq. 1a from eq. 12 gives

$$\varepsilon_s \Delta(v^{(m)} - \psi) = n^{(m)} - p^{(m)} - n + p = -(n+p)u^{(m)} \; , \tag{17}$$

using the approximation 14. It follows from 15, 16, 17 that the z_j are also eigenfunctions of the mapping $u^{(m)} \to v^{(m)} - \psi^{(m)}$, with corresponding eigenvalues $-(1+\lambda_j)$. Denoting by λ a bound on the λ_j, which are positive and accumulate only at zero, it follows (see ref. 7 for details) that appropriate values of α_m and β_m are given by

$$\text{a)} \quad \alpha_m = \frac{4}{\lambda} \frac{\cosh k\omega}{\cosh (k+1)\omega} \; , \; \beta_m = \frac{\cosh (k-1)\omega}{\cosh (k+1)\omega} \; ; \; \cosh \omega = 1 + 2/\lambda \; , \; k \neq 0; \tag{18}$$

$$\text{b)} \quad \alpha_m = 2/\lambda, \quad \beta_m = 0, \quad k = 0.$$

In 18 a, k is a nonzero integer depending on the iteration number m.

In each computation, k is set to zero initially and increased by one at each iteration. However, at any iteration for which $||v^{(m)}-\psi^{(m)}|| > ||v^{(m-1)}-\psi^{(m-1)}||$, the value of k is reset to zero, such a condition being considered an indication that nonlinear effects have become important.

It remains to estimate the constant λ; it follows from eq. 15 that λ is given by

$$\lambda = \sup_{z \in S} \frac{\int_{R_1} (n+p) z^2 dx\, dy}{\varepsilon_s ||z||_1^2} \quad . \tag{19}$$

The right side of (19) may be estimated by several methods. For example, it follows from Hölder's inequality and the Sobolev lemma that for any $q > 1$, there exists a constant $C(q)$ such that

$$\lambda \leq C(q) \left(\int_{R_1} (n+p)^q \, dx\, dy \right)^{1/q} \quad . \tag{20}$$

However, the sharply peaked form of the electron distribution function at the oxide interface makes estimates of this form too conservative to be of much practical value. Although eq. 20 is not strictly valid for $q = 1$, we have found empirically that a useful value of λ is given by

$$\lambda \simeq \frac{K}{4\varepsilon_s L} \int_{R_1} (n+p)\, dxdy \quad . \tag{21}$$

In evaluating eq. 21, it is also necessary to replace the unknown functions n,p with the initial approximations $n^{(o)}$, $p^{(o)}$.

4.4 Difference Approximations

The computations described above are readily performed on a computer by the use of finite difference approximations.

In view of the discontinuity of the dielectric constant at the oxide interface, it is convenient to discretize the Poisson equation in terms of the variables x,z, where

$$z=z(y)=\begin{cases} y/\varepsilon_{ox}, & y \le 0, \\ y/\varepsilon_{s}, & y \ge 0 ; \end{cases} \tag{22}$$

the Poisson equation then becomes

$$\begin{aligned} &\frac{\partial^2\psi}{\partial z^2} + \varepsilon_s^2 \frac{\partial^2\psi}{\partial x^2} = \varepsilon_s (N_A + p-n), \quad z > 0 ; \\ &\frac{\partial^2\psi}{\partial z^2} + \varepsilon_{ox}^2 \frac{\partial^2\psi}{\partial x^2} = 0 , \quad z < 0 ; \end{aligned} \tag{23}$$

since ψ and $\frac{\partial\psi}{\partial z}$ are continuous across the oxide interface (z=0), no additional relations there are necessary.

We superimpose a rectangular grid on the geometry of Fig. 60; the locations of the grid lines x_i, y_j are controlled by the following relations:

$$\begin{aligned} &\text{a)} \quad x_{i+1} - x_i \equiv h_x(i+1) = (1-\delta_x)h_x(i) \\ &\text{b)} \quad y_{j+1} - y_j \equiv h_y(j) = \varepsilon\, h_z(j) = \begin{cases} (1-\delta_y)h_y(j-1) & j < m \\ (1+\delta_y)h_y(j-1) & j > m+1 \end{cases} \\ &\text{c)} \quad h_z(m) = (1-\delta_y)h_z(m-1) = h_z(m+1)/(1+\delta_y) \\ &\text{d)} \quad \tilde{h}_x(i) \equiv 1/2(h_x(i)+h_x(i+1)), \quad \tilde{h}_y(j) \equiv 1/2(h_y(j-1)+h_y(j)) \end{aligned} \tag{24}$$

where m is the y-index corresponding to the oxide interface and the constants δ_x, δ_y are defined below.

The orientation of the discrete variables, relative to the grid points, is shown in Fig. 61, and the treatment of the boundary conditions in Figs. 62a and 62b. The orientation of the discrete variables near the oxide interface is shown in Fig. 62b, drawn approximately to scale for $\varepsilon_s = 3\varepsilon_{ox}$.

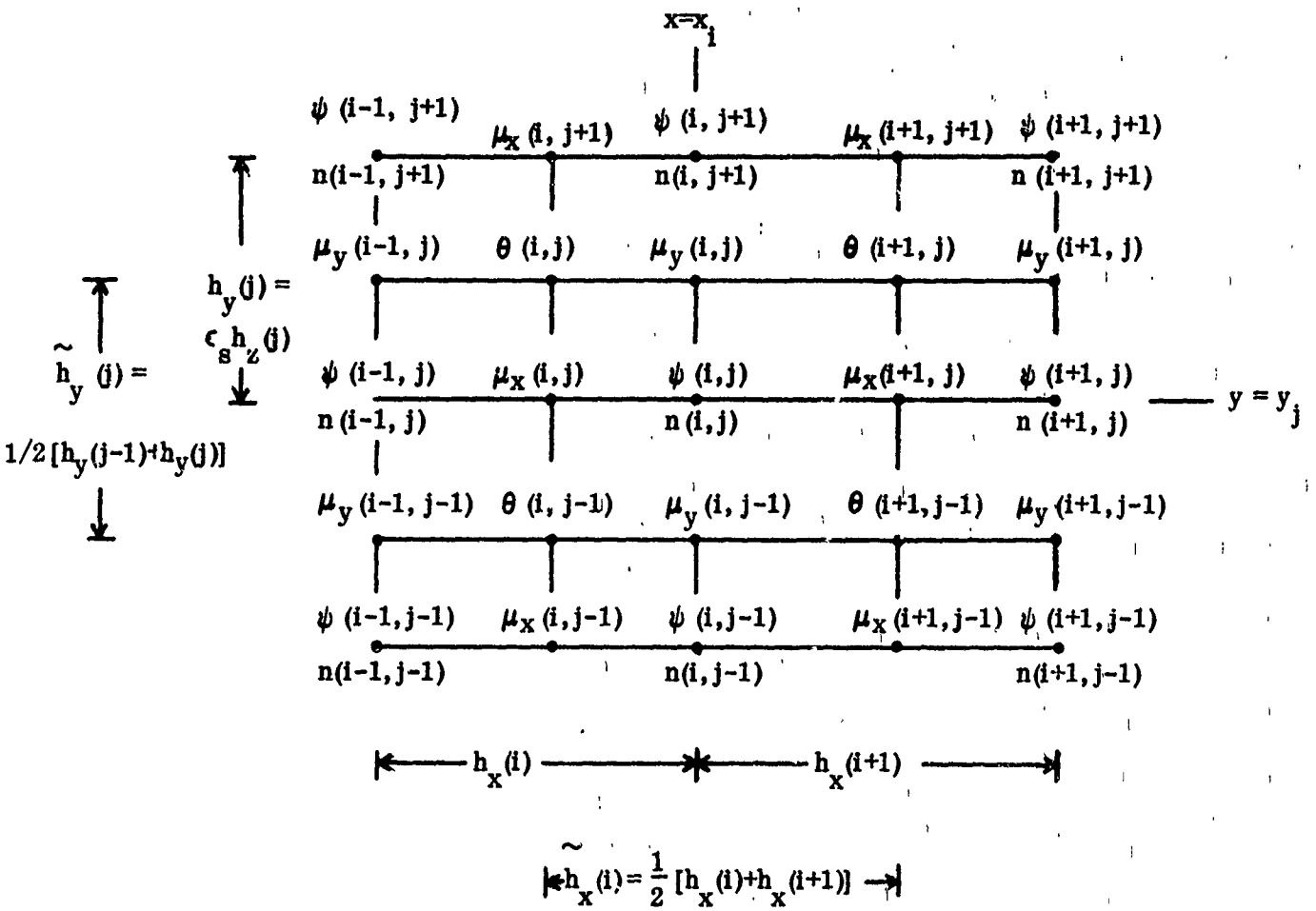

Fig. 61. Orientation of discrete variables.

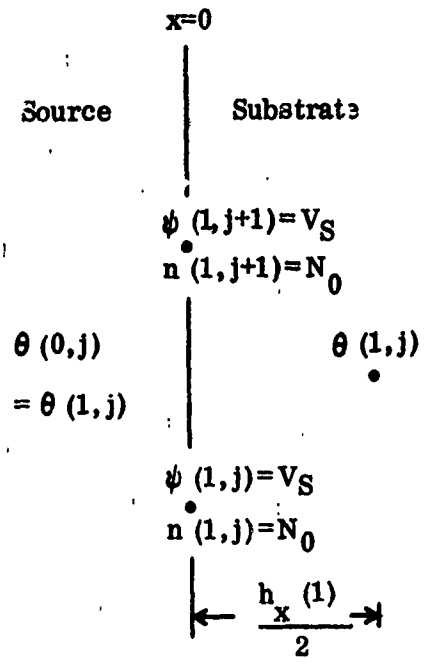

Fig. 62a. Boundary conditions at source contact.

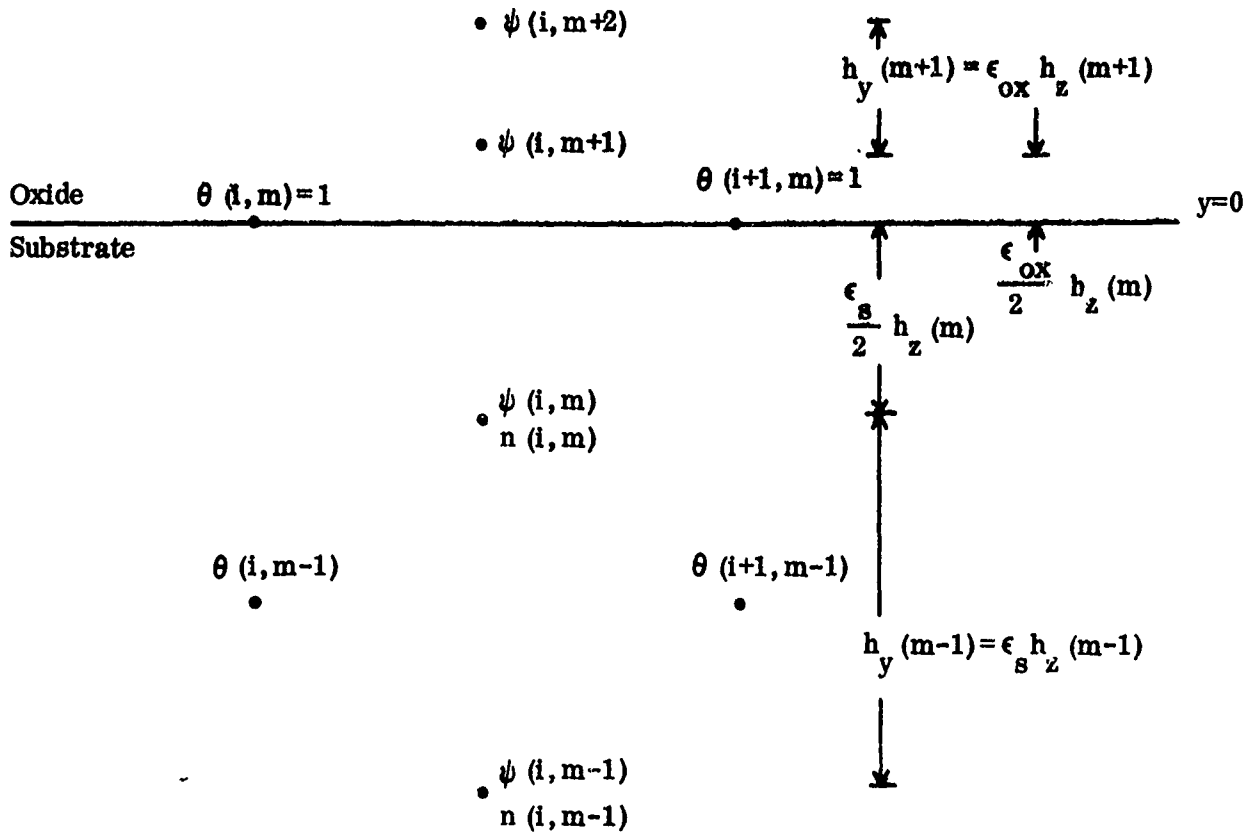

Fig. 62b. Boundary conditions at oxide interface.

By the standard "line integral" method, a difference approximation to eq. 23 is obtained whose local truncation error is of the order $h_x^2 + \delta_x h_x + h_z^2 + h_z \delta_y$. By the same method, we obtain difference approximations to the other required equations (8, 10, 11). Using a(i+1/2,j), a(i,j+1/2), respectively, as values of $e^{-\psi}$ between the pairs of points (i,j; i+1,j) and (i,j; i,j+1), we obtain the following:

$$h_y(j)\left(\frac{a(i,j+1/2)}{\mu_y(i,j)}\frac{(\theta(i+1,j)-\theta(i,j))}{\tilde{h}_x(i)} + \frac{a(i-1,j+1/2)}{\mu_y(i-1,j)}\frac{(\theta(i-1,j)-\theta(i,j))}{\tilde{h}_x(i-1)}\right) + \tag{25}$$

$$h_x(i)\left(\frac{a(i-1/2,j+1)}{\mu_x(i,j+1)}\frac{(\theta(i,j+1)-\theta(i,j))}{\tilde{h}_y(j+1)} + \frac{a(i-1/2,j)}{\mu_x(i,j)}\frac{(\theta(i,j-1)-\theta(i,j))}{\tilde{h}_y(j)}\right)=0;$$

$$J\sum_i h_x(i)\frac{a(i-1/2,m)}{\mu_x(i,m)}\frac{(1-\theta(i,m-1))}{\tilde{h}_y(m)} = 1-e^{-V_D}; \tag{26}$$

$$n(i,j)e^{-\psi(i,j)} - n(i+1,j)e^{-\psi(i+1,j)} = h_x(i+1)J\frac{a(i+1/2,j)}{\mu_x(i+1,j)}\frac{(\theta(i+1,j)-\theta(i+1,j-1))}{\tilde{h}_y(j)} \tag{27}$$

where the iteration denoting superscripts and the difference equation approximating to the Poisson equation have been omitted for simplicity.

The systems of linear equations generated by these difference approximations can all be efficiently solved. The coefficient matrix corresponding to the Poisson difference equations is symmetric and positive definite, and can readily be solved by the method of Dupont, Kendall, and Rachford.[10] To solve the system generated by eq. 25, we exploit the fact that the current flow in the IGFET device is essentially one-dimensional, i.e. the horizontal gradient terms in eq. 25 are relatively unimportant. We thus employ a line iteration,[11] solving successive tridiagonal systems of m equations generated by eq. 25 for a fixed i, and j varying. For this particular model, this iteration converges very rapidly. Equation 26 is explicit for J, and 27 for n; in practice, we solve eq. 27 for n(i+1,j) using the value of n(i,j) if $\psi(i+1,j) \leq \psi(i,j)$, and vice versa. This procedure suppresses the buildup of roundoff error; the consistency of the difference equations (25,26,27) insures the continuity of the obtained values of n(i,j).

4.5 A Critical Factor

It remains to define the quantities a(i+1/2,j), a(i,j+1/2) in terms of local values of the electrostatic potential; we employ the following approximation:

$$a(i+1/2,j) = \begin{cases} \dfrac{e^{-\psi(i+1,j)} - e^{-\psi(i,j)}}{\psi(i,j) - \psi(i+1,j)}, & \psi(i,j) \neq \psi(i+1,j) \\ e^{-\psi(i,j)}, & \psi(i,j) = \psi(i+1,j), \end{cases} \tag{28}$$

with an analogous expressing for a(1,j+1/2). This expression may be derived, as was done by Scharfetter and Gummel,[12] by noting that a(i+1,j) is identified with the average value of the function $\exp(-\psi)$ in the interval $x\varepsilon(x_i, x_{i+1})$, $y = y_j$, and assuming a linear dependence of ψ on x in this interval. This association becomes apparent in eq. 27, if one anticipates variations of several units of kT/q between $\psi(i,j)$ and $\psi(i+1,j)$; the left side of eq. 27 then cannot be regarded as an accurate difference approximation to $h_x \frac{\partial}{\partial x}(ne^{-\psi})$ in this interval. One is motivated to interpret eq. 27 as an integrated form of 7a between the points (x_i, y_j) and (x_{i+1}, y_j).

In one-dimensional models, it can also be shown by purely analytic arguments[13] that the expression (28) leads to a discretization error of $O(h^2)$ in the computed device current, without assuming small variations in ψ between successive grid points.

Finally, we show that an obvious alternative to eq. 28, which has been used in previous calculations on IGFET models,[2,3]

$$a^*(i+1/2,j) = \exp\left(-\frac{1}{2}(\psi(i,j) + \psi(i+1,j))\right), \tag{29}$$

can lead to serious error in the computed electron distribution.

Suppose (x_i, y_j) is a point near the drain junction at which a significant current density exists, in a device operating in its saturation region. We assume that

$$\psi(i+1,j) - \psi(i,j) \gg 1 \tag{30}$$

and consider the solution of eq. 27 for n(i,j). Setting

$$J(i+1/2,j)=J\,\frac{\theta(i+1,j)-\theta(i+1,j-1)}{h_y(j)}\,,\quad E(i+1/2,j)=\frac{\psi(i,j)-\psi(i+1,j)}{h_x(i+1)}$$

the local current density and electric field in the x-direction, respectively, we obtain from eqs. 27, 28, 30,

$$n(i,j)=\frac{J(i+1/2,j)}{\mu_x(i+1,j)E(i+1/2,j)}+O(\exp(\psi(i,j)-\psi(i+1,j)))\,; \tag{31}$$

the computed electron density is sufficient to support the local current density.

We can estimate the direction of the error in eq. 31. From the Poisson equation, it is clear that $\partial^2\psi/\partial x^2$ is positive in such an interval $x_i < x < x_{i+1}$; therefore the function $e^{-\psi}$ is larger in this interval than the linear approximation gives, and our expression for $a(i+1/2,j)$ is an underestimate. It follows that eq. 31 is an underestimate of the electron concentration.

If, however, eq. 29 is used instead of eq. 28, we obtain an electron concentration $n^*(i,j)$ given by

$$n^*(i,j)=n(i,j)\;\frac{\left(\frac{\psi(i+1,j)-\psi(i,j)}{2}\right)}{\sinh\left(\frac{\psi(i+1,j)-\psi(i,j)}{2}\right)} \tag{32}$$

where $n(i,j)$ is the electron concentration obtained in eq. 31. The value of $n^*(i,j)$ can thus be unrealistically small; for $\psi(i+1,j)-\psi(i,j) = 20$ (0.5 volt), we obtain $n(i,j)/n^*(i,j) \approx 10^3$.

4.6 Summary - Numerical Experiments

A unified presentation of the IGFET device characteristics as computed by the method described above, including a detailed comparison with existing theory appears elsewhere in this report. The empirical data given here is intended to support the numerical accuracy of the results presented.

Most of our computations have been run with 20 horizontal and 30 vertical grid points, with $\delta_x = 0.05$, $\delta_y = 0.1$ in eq, 24, Empirical accuracy estimates have been obtained by comparison of the results with results obtained using higher point densities and different values

of δ_x, δ_y; on this basis, the numerical (discretization) errors in the presented results are estimated at 5% of the quoted values of device currents, and 1% + 20 mV. of the quoted values of applied voltage along the substrate-oxide interface.

Numerical experiments have also been conducted to assess the importance of the difference approximation (28), the reliability of the formula (21), and the sensitivity of the computed results to the precise condition adopted for terminating a computation. The predictions of section 4.5 regarding the computed electron distribution have been observed in practice.

The use of a Chebyshev sequence of relaxation factors, as given by eq. 18, reduces the sensitivity of the number of iterations required in a computation to the sharpness of the adopted value of the constant λ and the stopping condition. The relation (21) has been found sufficiently conservative to insure convergence in all cases tested; in a few test cases, it was found to give a value of λ within a factor of two of the optimum.

The relaxed Picard iteration scheme described above converges essentially in the norm $||\psi-\psi^{(m)}||_1$;[5] previous analysis[13] suggests that in this case, the computed value of the device current will not depend strongly on the adopted terminating condition. Our experiments confirmed this, and the results given in this report were run using

$$||v^{(m)} - \psi^{(m)}|| \leq 10 \text{ mV.} \tag{33}$$

as the stopping criterion. We remark that this property is particular to the Picard iteration scheme.

4.7 Bipolar Transistor Model

A numerical analysis of the bipolar transistor has also been developed and incorporated into a computer program. In this section we present the associated equations and boundary conditions; the computation scheme is described in section 4.8.

The geometry associated with our approximation for the bipolar transistor is shown in Fig.63. Letting Ω denote the rectangle

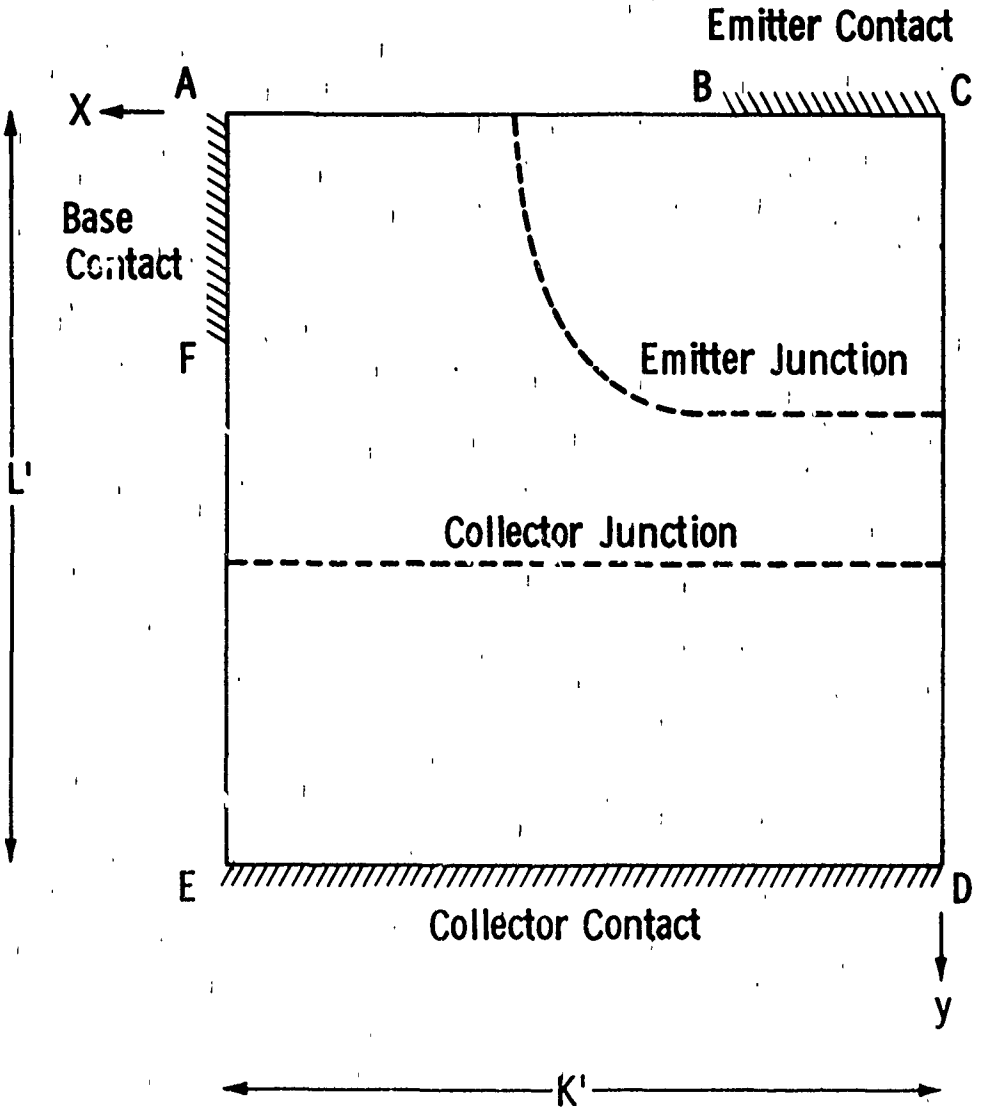

Fig. 63. Analytical model for a bipolar transistor.

$0 < x < K'$, $0 < y < L'$, we consider the electrostatic potential ψ and the local carrier densities n, p, to be described by the Poisson equation and continuity equations,

$$\varepsilon\Delta\psi + N(x,y) + p - n = 0, \tag{34}$$

$$\nabla\,(\mu_n e^{\psi}\nabla(e^{-\psi}n)) = R_n(n,p), \tag{35}$$

$$\nabla\,(\mu_p e^{-\psi}\nabla(e^{\psi}p)) = R_p(n,p); \quad (x,y) \text{ in } \Omega\ . \tag{36}$$

In eq. 34, ε is the dielectric constant and $N(x,y)$ the specified doping concentration, which determines the junction locations. In these calculations, we have used profiles of the form

$$N(x,y) = C_e f(x,y) - C_b e^{-ky^2} + C_c\ , \tag{37}$$

where C_e, C_b, C_c, k are constants and $f(x,y)$ is the solution of a two-dimensional emitter diffusion continuity equation. In the case of a constant impurity diffusion coefficient, a series solution to this problem is known;[14] alternatively, the numerical integration of this equation is not difficult. For simplicity, we have assumed in these calculations that the emitter contact and emitter diffusion mask opening coincide.

In eqs. 35, 36, we have assumed nondegenerate statistics; we also have used the standard Hall-Shockley-Read recombination term,

$$R_n(n,p) = R_p(n,p) = \frac{np - 1}{\tau_n(n+1) + \tau_p(p+1)}\ . \tag{38}$$

The electron and hole mobilities in eqs. 35, 36 are dependent on position, local electric field, and direction. Position dependency is included by the method of Caughey and Thomas.[15] In the collector junction space-charge region, where current flow is assisted by the local electric field, a correction to the carrier mobilities is applied by relations of the form

$$\mu_{n,x}(x,y) = \text{minimum}\left(\mu_{n,o}(x,y),\ \frac{\mu_{n,o}(x,y)V_{c,n}}{|E_x(x,y)|}^{1/2},\ V_{1,n}/|E_x(x,y)|\right) \tag{39}$$

where $\mu_{n,x}(x,y)$ is the x-direction electron mobility at the point (x,y), $\mu_{n,o}(x,y)$ is the position-dependent electron mobility at (x,y),

$E_x(x,y)$ the local x-direction component of the electric field, and $V_{c,n}$, $V_{l,n}$ are the "critical" and "limiting" electron velocities respectively.[16] Similar relations hold for $\mu_{n,y}(x,y)$, $\mu_{p,x}(x,y)$, $\mu_{p,y}(x,y)$, with the obvious changes in subscripts.

In the emitter junction space-charge region, current flow is opposed to the local electric field, and the mechanisms of carrier transport are not well understood. In this region, we have used relations of the form (eq. 39) with the "critical velocity" term omitted.

Boundary conditions for the system (eqs. 34-36) are formulated as follows: the contact boundary segments, AF, BC, DE in Fig. 63, are assumed to be infinite surface recombination velocity ohmic contacts, at which thermal equilibrium and charge neutrality apply; thus n, p, and ψ can all be specified. The applied bias voltages are defined in the usual way, in terms of the carrier quasi-Fermi potentials. Along the boundary segments AB, CD, EF, the normal component of the electron and hole current density is required to vanish. On the boundary segments AB and EF, the electrostatic potential distribution is computed from one-dimensional theory and specified, analogously to the IGFET model, eq. 5; on the segment CD, the normal component of the electric field is also required to vanish.

4.8 Computation Scheme for Bipolar Transistor Model

The computation scheme described in section 4.2 is inappropriate for the bipolar transistor model for two reasons: the finite recombination term precludes the stream function treatment of the carrier continuity equations, and the large carrier densities would reduce the admissible relaxation factors to very small values, as shown in eq. 21. We have adopted the following variation of Gummel's algorithm.[6]

At each point (x,y) in Ω, we assume that approximations to the electrostatic potential and the recombination term are known, denoted by $\psi^{(m)}(x,y)$, $R^{(m)}(x,y)$ respectively. We compute corresponding approximations to the carrier densities and "improved" approximations

to the potential and recombination terms by successively solving the following equations, each with its appropriate boundary conditions, as described in section 4.7.

$$\frac{\partial}{\partial x}\left(\mu_{n,x}\left(x,y,\frac{\partial\psi^{(m)}}{\partial x}\right)e^{\psi^{(m)}}\frac{\partial}{\partial x}\left(e^{-\psi^{(m)}}n^{(m)}\right)\right) + \frac{\partial}{\partial y}\left(\mu_{n,y}\left(x,y,\frac{\partial\psi^{(m)}}{\partial y}\right)e^{\psi^{(m)}}\frac{\partial}{\partial y}\left(e^{-\psi^{(m)}}n^{(m)}\right)\right) = R^{(m)} \tag{40}$$

$$\frac{\partial}{\partial x}\left(\mu_{p,x}\left(x,y,\frac{\partial\psi^{(m)}}{\partial x}\right)e^{-\psi^{(m)}}\frac{\partial}{\partial x}\left(e^{\psi^{(m)}}p^{(m)}\right)\right) + \frac{\partial}{\partial y}\left(\mu_{p,y}\left(x,y,\frac{\partial\psi^{(m)}}{\partial y}\right)e^{-\psi^{(m)}}\frac{\partial}{\partial y}\left(e^{\psi^{(m)}}p^{(m)}\right)\right) = R^{(m)} \tag{41}$$

$$R^{(m+1)} = (n^{(m)}p^{(m)} - 1) \,/\, (\tau_n(n^{(m)} + 1) + \tau_p(p^{(m)} + 1)) \tag{42}$$

$$\varepsilon\Delta(\psi^{(m+1)} - \psi^{(m)}) - \beta(n^{(m)} + p^{(m)})(\psi^{(m+1)} - \psi^{(m)}) + \alpha(\varepsilon\Delta\psi^{(m)} + N(x,y) + p^{(m)} - n^{(m)}) = 0 \tag{43}$$

In eq. 43, α, β are positive constants of order unity, suitable values of which have been determined empirically for each computation reported.[8]

Difference approximations to eqs. 40-43 are readily obtained, by methods similar to those used for the IGFET analysis, and described in section 4.4. The exponential average factors, described in section 4.5, have also been incorporated in the bipolar device computations.

The computations described in this report were run with 40 horizontal and 70 vertical grid points. The numerical accuracy of the

scheme was appraised by rerunning one computation, using a 75 x 120 point grid. The two values of collector current and h_{FE} obtained in this manner varied by 2% and 4%, respectively. Typical variation in the local carrier densities was 2%, and in the electrostatic potential distribution was 5 mV. We therefore anticipate numerical errors of this magnitude in the results presented.

4.9 References

1. W. Van Roosbroeck, *Bell Sys. Tech. J.*, 29, 560 (1954)
2. M. B. Barron, "Computer Aided Analysis of Insulated-Gate Field-Effect Transistors", *Stanford Electronics Lab Report* SEL-69-069 (1969)
3. D. Vandorpe and N. H. Xuong, *Electronics Letters*, 7, 47 (1971)
4. C. T. Sah, *IEEE Trans. Elec. Dev.*, ED-11, 324 (1964)
5. M. S. Mock, "On Equations Describing Steady-State Carrier Distributions in a Semiconductor Device" (to be published)
6. H. K. Gummel, *IEEE Trans. Elec. Dev.*, ED-11, 455 (1964)
7. G. E. Forsythe and W. R. Wasow, "*Finite-Difference Methods for Partial Differential Equations*", pp. 226-235, Wiley, New York (1960)
8. M. S. Mock, "On the Convergence of Gummel's Numerical Algorithm", *Solid-State Electronics*, 15, 1 (1972)
9. G. Hellwig, "*Partial Differential Equations*", pp. 222, Blaisdell (1964)
10. T. Dupont, R. P. Kendall and H. H. Rachford, Jr., *SIAM J. Num. Anal.*, 5, 559 (1968)
11. R. S. Varga, "*Matrix Iterative Analysis*", pp. 194, Prentice-Hall (1962)
12. D. L. Scharfetter and H. K. Gummel, *IEEE Trans. Elec. Dev.*, 16, 64 (1969)
13. M. S. Mock, "On the Computation of Semiconductor Device Current Characteristics by Finite Difference Methods", (to be published)
14. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. Dev.*, 9, 179 (1965)
15. D. M. Caughey and R. F. Thomas, *Proc. IEEE (Letters)*, 55,2192 (1967)
16. E. J. Ryder, *Phys. Rev.*, 90, 766 (1953)